ABSTRACT

Dissertation: EXAMINING EVIDENCE OF RELIABILITY
                                      AND VALIDITY OF MENTAL HEALTH
                                      INDICATORS ON A REVISED NATIONAL
                                      SURVEY MEASURING COLLEGE
                                      STUDENT HEALTH

                                      Theresa K. Jackson, Doctor of Philosophy, 2008

Directed By: Elbert D. Glover, Ph.D., Professor and Chair
                                      Department of Public and Community Health
                                      University of Maryland School of Public Health

The American College Health Association-National College Health Assessment

(ACHA-NCHA), which began surveying college students in 2001, is currently the only

large scale survey available for colleges and universities to measure a variety of health

constructs among their student populations. Beginning in December 2005, the survey

underwent an extensive revisions process in an effort to improve its measurement quality

and to better capture the current health status of college students. Revisions were guided

by changing student health priorities, feedback from respondents, literature focused on

characteristics of reliable and valid survey questions, and the Model of Survey Response.

As mental health concerns continue to rise on campuses today, this content area of the

ACHA-NCHA was dramatically expanded, as it now includes constructs such as

diagnosis and treatment with multiple mental health conditions, difficult life experiences,

stress, and help-seeking. This dissertation, which is a secondary data analysis of data

collected from the original and revised ACHA-NCHA surveys during an experimental

field pre-test of the modified survey, documents the survey revisions process, provides

results from more than 40 mental health indicators by various demographic

characteristics, and establishes the reliability and validity of the mental health indicators. The data analyzed in this study were collected from students from 7 U.S. college and universities who were randomized to complete either the original (final N = 6,216) or the revised (final N = 6,110) online ACHA-NCHA from February through May 2007. It was hypothesized that (1) changes to survey indicators designed to measure comparable constructs would result in significant differences in student response patterns across versions of the ACHA-NCHA; (2) mental health indicators on the revised survey would demonstrate evidence of internal consistency reliability, construct-related validity, and criterion-related validity; and (3) modified mental health indicators on the revised ACHA-NCHA would demonstrate greater evidence of reliability and validity than comparable indicators on the original ACHA-NCHA. Findings from this study at least partially support all hypotheses, and the revised ACHA-NCHA demonstrates preliminary evidence that is it a psychometrically sound survey tool to measure college student mental health constructs.

EXAMINING EVIDENCE OF RELIABILITY AND VALIDITY OF MENTAL
HEALTH INDICATORS ON A REVISED NATIONAL SURVEY
MEASURING COLLEGE STUDENT HEALTH

By

Theresa Katherine Jackson

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2008

Advisory Committee:
Elbert D. Glover, Ph.D., Professor, Chair
Mary Ann Hoffman, Ph.D., Professor, Dean's Representative
Robin Sawyer, Ph.D., Associate Professor
Katherine Sharp, Ph.D., Instructor
Min Chi Wang, Ph.D., Professor

PREFACE

In May 2005, only a month before I intended to embark upon my journey as a doctoral student in the Department of Public and Community Health at the University of Maryland at College Park, I emailed the Research Director at the American College Health Association (ACHA), E. Victor Leino, Ph.D., about the possibility of working part-time at ACHA during my studies. Victor forwarded my email to Mary Hoban, Ph.D., who ultimately arranged my interview. Little did I know that that email would spur a sequence of events that would ultimately lead to the most profound research—and professional—experience of my life to date.

When I began working at ACHA as a Research Assistant on the ACHA-National College Health Assessment (ACHA-NCHA), I organized paperwork, contacted institutions to answer questions, and assisted Victor—who was, at the time, the single person behind the entire ACHA-NCHA survey operation—with tasks when necessary. Ultimately, my hours per week and responsibilities expanded, the ACHA-NCHA continued to grow exponentially as more campuses chose to participate in data collection efforts, and Mary, Victor, and I became the ACHA-NCHA team.

In December 2005, under Mary and Victor's lead, the ACHA-NCHA underwent an extensive revisions process, the numerous steps for which are described in the many pages of this dissertation. I was extremely fortunate to participate in nearly every step along the way and my responsibilities in the process included compiling and theming a list of potential items to add to the survey, participating in all meetings of the revisions committee, serving as the only student member of the committee, documenting comments made by the eight other brilliant members of the committee for nearly two years' worth of work, assisting with the programming of all survey efforts for the revised and original survey web implementation in Spring 2007, programming the syntax for downloading data for the revised survey, downloading all data from the revised survey pre-test, and analyzing a variety of the results from the revised survey. This dissertation expands upon my work at the ACHA and focuses on only one small subset of survey items, but it also documents many of the steps by the entire ACHA-NCHA revisions

committee to create and evaluate what I believe is, without question, one of the most important survey instruments in use on college campuses today.

To have been a part of this process has been a truly fulfilling experience and one I will carry with me for the rest of my life as a doctoral-level researcher. I learned more than I ever thought possible about the details and minutia associated with survey research, downloading and analyzing data, and working with the leaders in the field about which I feel most passionate—college student health. I sincerely thank you for taking the time to read this research, and I hope you find it as interesting and important as I continue to.

Theresa K. Jackson
Department of Public and Community Health
University of Maryland School of Public Health

ACKNOWLEDGEMENTS

The completion of this research, and ultimately my degree, would not have been possible without the support, guidance, and direction of so many influential people in my life. Therefore, it is with tremendous respect, gratitude, and appreciation that I would like to acknowledge the following individuals for their assistance throughout this process.

Mom and Dad

Though we all know I have always been and continue to remain a "chatty" child, I truly do not have enough words of thanks for everything you have supported me in accomplishing. Your unwavering, unconditional support has been the one constant I cannot remember ever living without. Thank you for always believing in me, for encouraging me to follow the path I felt was right for me, and for always providing a safe place and respite when I needed it. And thank you for not only dreaming big for me but for allowing me to dream big for myself. You truly are the most amazing parents and role models, and I hope that one day I will be able to support my children using the example you set for me. I love you and I certainly could not have done this without you.

Patrick and Erin

There are people in our lives who ultimately just make everything seem like it is always going to be okay and around whom we don't feel the need to be anything but ourselves. For me, you are two of those people. Thank you for always putting up with my antics, for supporting me through the roughest parts of the last few years, and for making me a better person just by knowing you. That you are my family is the best gift I could ever think of receiving.

The Lambden/Jackson clan

Oh, where to begin. You each continue to be such an influential and important source of love and guidance in my life. No matter where I have gone or what I have done, you are my home. Thank you for keeping me grounded, for always making me laugh, for providing me with the strongest of roots, and for being the people with whom I best identify. I love you all so much.

Dr. Glover

I have to first thank you for being such a wonderful mentor and advisor. I can honestly say that I always left your office feeling better and more empowered than when I entered. You have taught me so much, and I will always be grateful for your advice, wisdom, guidance, and assistance. I feel honored to be your first doctoral advisee at the University of Maryland. Thank you for helping me to grow as a student, a researcher, and a person over the last three years.

Drs. Sawyer, Sharp, Wang, and Hoffman

Thank you very much for serving on my dissertation committee and for adding your insight and expertise to this work. Despite each of your tremendously busy schedules, you have contributed so much and I will always be grateful to you for your assistance with this project. I honestly believe that I have the best dissertation committee of any student at the University.

Mary and Victor

From my heart, thank you both for serving as my most influential professional mentors. You will never know the appreciation I feel for the experiences I had while working at the American College Health Association and with the ACHA-NCHA. Thank you for taking the chance to hire me in 2005 and for putting up with my unending questions, quippy remarks ("I'm saying!"), and pathological peppiness. I learned so much from each of you in terms of my knowledge and professionalism. If I have the opportunity to supervise others in the future, I can only hope to provide the type of supportive atmosphere that you created for me during the last three years.

My friends

To all of you, but particularly to Kristin, Jamie, Jill, Ali, Chris, JJ, Dave, Jessie, Angeli, and James: Thank you for always being interested in my never-ending ramblings about this project (well, and also to my never-ending ramblings about life in general). You are each so important to me and have served as an amazing source of support in my life. Plus, because of you I have saved thousands of dollars in therapy. So thanks for that. ☺

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I: INTRODUCTION

Introduction

Currently, nearly half of students who graduate high school in the United States (42 percent of males, 46 percent of females) enroll in a post-secondary institution (PSI) for higher education (Baum & Payea, 2005), the most common form of which are colleges and universities. PSIs are defined as organizations open to the public that have as their primary mission the provision of postsecondary education (Knapp, Kelly-Reid, Whitmore, & Miller, 2007). Postsecondary education encompasses formal instructional programs with a curriculum designed primarily for students who are beyond the compulsory age for high school, which includes academic, vocational, and continuing professional education programs and excludes institutions that offer only avocational (leisure) and adult basic education programs. In 2005, 17.9 million students enrolled in more than 6,500 PSIs (Knapp et al., 2007).

The benefits of a college education to both individuals and society have been well-documented in the literature (e.g., Institute for Higher Education Policy, 1998; Weale, 1992; McMahon, 1999; Baum & Payea, 2005; Rowley & Hurtado, 2002). For example, a recent study by the College Board suggests that at the individual level, college graduates earn about 73 percent more than typical high school graduates, experience poverty at one-third the rate of high school graduates, and are more likely to perceive themselves as being in good health and less likely to smoke than high school graduates (Baum & Payea, 2005). Non-monetary individual benefits of higher education include the tendency for postsecondary students to become more open-minded, more cultured, more

rational, more consistent and less authoritarian (Rowley & Hurtado, 2002). From a societal standpoint, adults with higher levels of education are more likely to engage in various civically-minded activities, such as organized volunteer work, voting, and blood donation. Further, the rate of incarceration of adults with some college education is only one-quarter of that for high school graduates (Baum & Payea). Additionally, college attendance has been shown to "decrease prejudice, enhance knowledge of world affairs and enhance social status" while increasing economic and job security for those who earn bachelor's degrees (Rowley & Hurtado).

As such, it is not surprising that educational attainment remains a top concern of families, politicians, and the general public. Students, their families, and tax-payers spend billions of dollars each year to finance college educations, and costs of education continue to rise (The College Board, 2007). The average total charges per year per student in 2007-2008 (including tuition, fees, room and board) are $32,307, up 5.9 percent from 2006-2007. Further, state and local appropriations per student were $6,695 in 2005-2006. Despite these rising financial costs, enrollment in PSIs continues to grow, as students and their families recognize the benefits higher education (The College Board).

Yet, despite the numerous benefits of higher education, college students today are at risk for a variety of health concerns, including high-risk alcohol use, sexually transmitted infections, depression, suicide, and accidents (e.g., Douglas, Collins, Warren et al., 1997; Centers for Disease Control & Prevention [CDC], 2007). Grace (1997) states,

> "Young adults' health needs have been regularly overlooked by public health professionals and legislators because the stereotype of a typical college student is inaccurate. Once depicted as a robust and healthy group, students today are older and

have more serious, chronic, and sometimes catastrophic medical problems. More important, an epidemic of long-standing unhealthy behaviors is occurring on campuses today. College students are not adopting the preventive behaviors they know about. In spite of our best efforts at health education, these risky behavior patterns continue while the health and economic consequences are postponed to future generations. National health indicators suggest that individuals in this age category are experiencing the least overall reduction in morbidity and mortality rates" (*Lessons from the Past,* ¶3).

These health concerns have a significant impact on academic performance (American College Health Association [ACHA], 2007a), can reduce the likelihood of degree completion (Kessler, Foster, Saunders, & Stang, 1995), and can create behavioral and health problems that follow these students into their adult lives (e.g., Sparling & Snow, 2002; McCarty, Ebel, Garrison, et al., 2004). Therefore, an adequate understanding of these issues and the means through which we can best intervene is a critical concern of college and university administrators, faculty, and staff as well as public health professionals (Grace, 1997; ACHA, 2007a; CDC, 1997).

Accurate measurement of student health constructs, including beliefs, behaviors, and experiences, is imperative to setting priorities, monitoring trends, demonstrating need for intervention, and evaluating program effectiveness (e.g., Fowler, 1995; de Vaus, 1995; CDC, 1997; Di Iorio, 2005; ACHA, 2007a; Presley, Meilman, & Lyerla, 1994). As is the case in most fields within the social sciences (Marsh, 1982; Fowler, 2002; de Vaus, 1995), college health relies heavily on surveys as the primary mechanism through which data are collected in this population. The general principles of survey research, as delineated by Singleton & Straits (1999, p. 239) are as follows:

1. A large number of respondents are chosen through probability sampling procedures to represent the population of interest.
2. Systematic questionnaire or interview procedures are used to ask prescribed questions of respondents and record their answers.
3. Answers are numerically coded and analyzed with the aid of statistical software.

Surveys, or questionnaires designed to address specific research questions (Di Iorio, 2005), vary greatly in their format, length, and scope of measurement. Despite the wide range of health concerns experienced by college students, nearly every survey tool used on campuses today focuses solely on one aspect of student health (e.g., alcohol and drug use, sexual health). To date, there has only been one nationally generalizable survey study of comprehensive student health, the CDC's National Collegiate Health Risk Behavior Survey (NCHRBS), which was implemented in 1995 (Douglas et al., 1997; CDC, 1997).

Since Spring 2000, the American College Health Association's (please refer to Appendix A for a brief overview of the association) ACHA-National College Health Assessment (ACHA-NCHA), a comprehensive self-report survey collecting data on nearly 300 items, has served as an invaluable tool measuring college student health (ACHA, 2007b). An assessment endeavor that stemmed from the NCHRBS, the ACHA-NCHA is comprised of six sections: (1) health, health education, and safety; (2) alcohol, tobacco, and other drug use; (3) sexual health; (4) mental and physical health; (5) impediments to academic performance; and (6) demographic characteristics (ACHA, 2007b).

To date, nearly 400 college campuses across the United States have self-selected to participate in data collection efforts (see Table 1.1 for a comprehensive picture of survey participation since 2000). Data from the ACHA-NCHA, which have been collected from more than 450,000 students on campuses in all 50 states as well as Canada since its inception (ACHA, 2007b), have been and continue to be used for a multitude of purposes. These data, which are collected each fall and spring, are cited in a variety of

sources including academic literature, college-wide publications, social norms

campaigns, and mass media (ACHA, 2007b). Further, each year beginning with Spring

2003, ACHA publishes the results of its spring survey period in the *Journal of American

College Health* as a current report of student health (ACHA, 2005, 2006a, 2006b, 2007a).

As such, reliable and valid data collection is of the utmost priority to the American

College Health Association, its members, and survey administrators, as well as the

stakeholders who rely on these data for their decision making, assessment, and evaluation

purposes (ACHA, 2004a, 2007a, 2007b).

Table 1.1. Participation in the ACHA-NCHA from Spring 2000 through Spring 2007*.

ACHA-NCHA Survey Periods	Number of participating PSIs	Total sample (N)
Spring 2000	28	16,024
Fall 2000	20	10,413
Spring 2001	31	16,813
Fall 2001	8	4,717
Spring 2002	44	28,258
Fall 2002	20	10,374
Spring 2003	33	19,497
Fall 2003	21	11,990
Spring 2004	74	47,202
Fall 2004	50	24,804
Spring 2005	71	54,111
Fall 2005	29	16,832
Spring 2006	117	94,806
Fall 2006	34	23,863
Spring 2007	107	71,860
Cumulative Participation	**687****	**451,564**

Note: Participation rates are provided for the original version of the ACHA-NCHA only. Additionally, participation numbers include only PSIs that utilized census or random sampling techniques. Institutions that utilize non-random sampling techniques are not included in totals nor in reference group summaries generated by ACHA.

**This total number does not represent unique PSIs. Numerous PSIs participated in the NCHA in more than one survey administration period.

As a result of feedback from student participants, college health professionals, public health experts, and other national organizations; changing assessment needs and priorities; and developments in survey literature, beginning in fall 2005, the ACHA-NCHA underwent an extensive revision process. This researcher served as a Research Assistant at the ACHA from July 2005 through December 2007 and was hence integrally involved in the modification and expansion of the ACHA-NCHA to its revised form. During this process, each of the six sections was dramatically revised with hopes that the revised instrument would become more psychometrically sound and would better meet the assessment needs of PSIs, their students and their administrations (see Appendices B.1 and B.2 for full versions of the original and revised ACHA-NCHA survey instruments).

This iterative revision process utilized several strategies described in survey development literature (e.g., Presser, Couper, Lesser, et al., 2004; Fowler, 1995; Trochim, 2006; DeVellis, 2003; Singlteon & Straits, 1999). It was guided in part by current studies on what constitutes reliable and valid survey questions and was further informed by a recently developed theory, Tourangeau's (1984) Model of Survey Response. A brief overview of the revisions process is described below:

- In September 2005, this researcher generated a pool of potential new survey items as a result of a literature review, examination of smaller scale surveys, and consultation of other national instruments measuring aspects of college, adolescent, and adult health;

- In November 2005, ACHA provided its membership with a two-week comment period, during which all members were encouraged to provide feedback on the ACHA-NCHA, the changes they hoped would be made, and their experiences with implementation on their campuses;

- In December 2005, a panel of 8 experts and 3 ACHA staff members (including this researcher) was initially convened to evaluate the applicability, appropriateness, usefulness, and psychometric properties of each item on the original ACHA-NCHA (see Appendix C for a list of experts and their positions). Each expert individually evaluated whether individual items on the instrument should be kept as is, kept with modifications, or eliminated. Results were compiled and patterns were evaluated. This panel of experts and ACHA staff members then became the ACHA-NCHA revisions committee;

- This researcher compiled and themed feedback from each of the three steps above. Themes were then used to generate revision priorities;

- From January 2006 through December 2006, in a constant and iterative process, the revisions committee examined each of the six ACHA-NCHA sections item by item. Each item was reviewed and decisions were made as to what revisions, if any, needed to be made. Common revisions included (1) stem re-wording, (2) clarification of response instructions, (3) modification of response options, and (4) modification of recall time period. Items deemed psychometrically unsound or no longer useful were removed from the survey, and new items were added when necessary. The committee required consensus from all members before any modifications, deletions, or additions were finalized;

- In January 2007, ACHA staff, including this researcher, programmed the revised instrument into an online survey application and prepared it for field pre-testing among the target population (Singleton & Straits, 1999). Simultaneously, ACHA staff, including this researcher, recruited seven self-selected PSIs to participate in the initial pre-test of the revised ACHA-NCHA (see Chapter Three for a more comprehensive discussion of pre-testing methodology);

- From February to May 2007, students from each of the seven PSIs participated in the field pre-test of data collection. ACHA researchers randomized students from each school to complete either the original version of the ACHA-NHCA or the revised version so results from the two versions could be directly compared (Presser et al., 2004). ACHA staff members, including this researcher, downloaded and compiled results;

- In April 2007, two ACHA staff members, including this researcher, conducted focus groups of students (n=20) who participated in the pilot to better understand their experiences and interpretation of items;

- From May 2007 through present, each item on the revised survey is being psychometrically evaluated using response analysis, a field pre-test technique in which the responses of pre-test respondents are tabulated and examined for problematic response patterns (Singleton & Straits, 1999). Once results are finalized, the survey will undergo additional revisions and further pilot testing with hopes for final instrument implementation and adoption in Fall 2008 or Spring 2009.

Perhaps no one section changed as dramatically as that which measures student mental and physical health. Although college student mental health has been a prominent concern for decades, recent events such as the Virginia Tech massacre, Northern Illinois campus shootings, and high-profile student suicides have highlighted the important mental health crisis on campuses nationwide (Shea, 2002; Cook, 2007; Voelker, 2007). In the time since the ACHA-NCHA was first introduced, mental health concerns of college students, including depression, suicide, self-harm, anxiety, stress, eating disorders, help-seeking, and intervention, have been researched and addressed extensively (e.g., Suicide Prevention Resource Center [SPRC], 2004; Leino & Kisch, 2005; Kadison, 2006;

Kitzrow, 2003; Soet & Sevig, 2006). Thus, the need for more comprehensive and psychometrically sound mental health indicators on the ACHA-NCHA was a consistent theme that emerged throughout the aforementioned revisions process.

Although the original ACHA-NCHA contained a series of items assessing mental health, data from which have been cited in dozens of academic and media publications (e.g., SPRC, 2004; Leino & Kisch, 2005; Kisch, Leino, & Sliverman, 2005; Blom & Beckley, 2005; Kadison & DiGeronimo, 2004), the revised ACHA-NCHA contains numerous additional mental health items as well as modifications to indicators that were on the original instrument, such as the negative affect scale and questions asking about various mental health diagnoses. Chapter Three provides an extensive overview of these additions and modifications.

This dissertation, which is part of a larger study preliminarily validating the complete revised ACHA-NCHA, is an examination of the psychometric properties of the expanded and modified mental health items and indicators on the survey. It was the intent of this researcher—who participated fully in all aspects of the survey revisions process—as well the entire ACHA-NCHA revisions committee to utilize current practices in survey research methodology, recommendations of what constitutes reliable and valid survey questions, and the Model of Survey Response to understand how to best assess an array of student health constructs. Thus, it is expected that the revised ACHA-NCHA will be a more psychometrically sound, comprehensive, and accurate survey instrument tool that will be used to collect information from students about a variety of health issues for years to come. Hence, the purpose of this study is two-fold: (1) to examine how modifications to the items on the original ACHA-NCHA change patterns of student responses to these

items; and (2) to examine evidence of preliminary reliability and validity of the revised and new mental health indicators on this large-scale survey in a pre-tested sample.

Statement of the Problem

According to the National Co-morbidity Survey (Kessler, 1994; Kessler, McGonagle, Zhao, et al., 1994) and its more recent replication (Kessler et al, 2005), nearly one in two U.S. adults will experience a mental disorder—defined as "health conditions that are characterized by alterations in thinking, mood, or behavior (or some combination thereof) associated with distress and/or impaired functioning" (U.S. Department of Health & Human Services, [USDHHS], 1999)—in their lifetime, the majority of which are stress-related anxiety, impulse control, mood, or substance abuse disorders. Further, the National Institute of Mental Health (NIMH, 2007a) claims approximately 26.2 percent of Americans ages 18 and older, or nearly 58 million U.S. residents, suffer from a diagnosable mental disorder in a given year. Many people suffer from more than one mental disorder at a given time, and nearly half (45 percent) of those with any mental disorder meet criteria for 2 or more disorders, with disorder severity strongly related to comorbidity (NIMH, 2007a).

Mental disorders are the leading cause of disability in the U.S. and Canada for ages 15-44 (NIMH, 2007a) and a recent study found that 1.3 billion disability days resulted from mental conditions in the U.S. over one calendar year (Merikangas, Ames, Cui, et al., 2007). Further, mental illness, including suicide, accounts for over 15 percent of the burden of disease in established market economies, such as the United States, which is more than the disease burden caused by all cancers (NIMH, 2007a).

Adolescence is a dynamic period of life transition during which many mental health issues first appear (e.g., National Mental Health Association & The Jed Foundation [NMHA & Jed], 2002; Blumenthal & Kepfer, 1990; Kessler et al., 2005), and the prevalence of mental health disorders in this age group has risen dramatically over the past few decades (Shea, 2002; Whitaker, 1989; Gallagher, 2007; Benton, Robertson, Tseng, et al., 2003). Kessler and colleagues (2005) found that 52.4 percent of 18-29 year-olds experienced a mental illness, and epidemiological studies indicate that the past-year prevalence rate of mental illness is highest (39 percent) for youth in the age category of 15–21 years old (Mowbray, Megivern, Mandiberg, et al., 2006). As this is the age at which many individuals first begin higher education, it is not surprising that numerous studies point to serious mental health problems among college students (e.g., SPRC, 2004; Kisch & Leino, 2005; Collins & Mowbray, 2005; Benton et al., 2003).

Common mental health problems reported in college populations include depression, anxiety, eating disorders, alcohol and other substance abuse problems, suicide, self-mutilation, and various other self-destructive and reckless behaviors (Gallagher, 2007; Benton et al., 2003; Cook, 2007; Kisch & Leino, 2005; Kadison & DiGeronimo, 2004; Sharkin, 2006). It has been estimated that approximately 12-18 percent of students on college campuses have a diagnosable mental illness (Mowbray et al., 2006), and much higher percentages report exceedingly high stress levels and mental health impairment (Oswalt & Finkelberg, 1995; Grace, 1997; Ross, Niebling, & Heckert, 1999; Cook, 2007).

A recent study found the number of freshmen reporting less than average emotional health has been steadily rising since 1985 (Sax, Lindholm, Astin, et al., 2001),

and a survey of 292 PSI counseling center directors found that 91.5 percent believed the recent trend toward greater number of students with severe psychological problems continues to be true on their campuses (Gallagher, 2007). Similarly, a study of student affairs administrators reported they were spending more time dealing with troubled students and had seen marked increases in the following serious mental health problems on campus: Eating disorders (+58 percent), Drug abuse (+42 percent), Alcohol abuse (+35 percent), Classroom disruption (+44 percent), Gambling (+25 percent), and Suicide attempts (+23 percent) (Levine & Cureton, 1998a).

Mental illness affects a student's motivation, concentration, and ability to appropriately engage in social interactions, all of which are critical factors in determining academic success (Padron, 2006). Fortunately, there are strong indications that students with psychiatric diagnoses can and do succeed in college if they receive appropriate treatment and supports (Collins & Mowbray, 2005). However, if left unrecognized and untreated, mental health problems may lead to students dropping out or failing out of college (Cook, 2007), and once educational attainments are interrupted, they are not likely to be regained; for most students, this begins a trajectory of poor vocational outcomes and poverty (Kessler et al., 1995). Additionally, mental health problems among college students are associated with suicide attempts and ideation as well as engagement in other risky, dangerous behaviors that may result in serious injury, disability, or death (Cook, 2007). Untreated mental illnesses—specifically depression, bipolar disorder, schizophrenia, and substance abuse—are the leading contributory causes of suicide in young adults (Goldsmith, Pellmar, Kleinman, & Bunney, 2002). At an estimated rate of 7.5 per 100,000 (Silverman et al., 1997), suicide is the second leading cause of death

among college students, surpassed only by accidents (NMHA & Jed, 2002). It is

estimated that as many as 1 in 12 students have a suicide plan and more than 1,000

students take their own lives each year (NMHA & Jed).

Given the increasing prevalence of mental disorders on college campuses, the fact

that mental health is integrally intertwined with all aspects of health (USDHHS, 1999),

and the devastating consequences of untreated and unmanaged mental illness, it is not

surprising there has been an urgent call for prevention, detection, and treatment of mental

disorders on college campuses across the country (Voelker, 2007; SPRC, 2004; NMHA

& Jed, 2002). Kitzrow (2003) states, "It is important for administrators, faculty, and staff

to understand the profound impact that mental health problems can have on all aspects of

campus life, and to treat mental health issues as an institutional responsibility and

priority" (p. 178). The literature also demonstrates a need to better understand the risk

and protective factors for mental health—those factors or characteristics that, when

present, serve to increase or reduce the likelihood that a person will develop a mental

disorder—and the mechanisms through which they operate (USDHHS, 1999).

Sound research and data collection are an important aspect of this understanding.

As is the case for numerous social science phenomena, survey data are a prominent

source of information for college student mental health issues (e.g., Pealer, Weiler, Pigg,

Miller, & Dorman, 2001; NMHA & Jed, 2002; SPRC, 2004). These data, which are

collected from administrators, counseling center directors, student affairs professionals,

faculty, staff, and students, provide a wealth of information in regards to current trends in

mental health, as well student health in general. Data can be used for an extensive host of

purposes, including the focusing of prevention efforts, prioritization of the allocation of

sparse resources, mobilization of community support for prevention and intervention,

correction of misperceptions about student health, establishment of quantitative support

for a request for funding, and exploration of complex relationships between mental health

and other health and demographic constructs.

The most common types of measures used to collect data about health and health

behaviors are survey instruments in which participants provide self report of knowledge,

attitudes, intentions, behaviors, and experiences (Di Iorio, 2005). Though the literature

indicates that a wide array of student self-report instruments exist to measure college

student health, most are limited in their scope and/or reach. For example, numerous

national large-scale surveys exist to monitor students' alcohol and other drug use (e.g.,

CORE survey, Monitoring the Future, College Alcohol Study). Some of these

instruments contain items that ask about very limited student mental health concerns

(e.g., depression), but they capture only a very small portion of the overall picture of

student health.

The need for a national survey measuring a wide variety of aspects of student

health has been well-documented (CDC, 1997; Douglas et al., 1997; ACHA, 2007a). To

date, the NCHRBS is the only large-scale, comprehensive, nationally generalizable study

of college student health (CDC, 1997; Douglas et al., 1997). This survey, the results of

which have been analyzed in dozens and cited in hundreds of publications, covered six

aspects of student risk behaviors: (1) behaviors that contribute to unintentional and

intentional injuries; (2) tobacco use; (3) alcohol and other drug use; (4) sexual behaviors;

(5) unhealthy dietary behaviors; and (6) physical inactivity. Yet, though it is more

comprehensive than most surveys measuring student health, the NCHRBS failed to address mental health among college students.

As such, the ACHA-NCHA is currently the only large-scale survey effort that measures such a wide array of student health issues (Soet & Sevig, 2006). Not surprisingly, data from the ACHA-NCHA—and its mental health indicators in particular—have been cited in a multitude of publications and are used for an extensive variety of purposes. For the past seven years, the ACHA-NCHA has been viewed as one of the primary sources of college student mental health data (e.g., NMHA & Jed, 2002; SPRC, 2004; Soet & Sevig, 2006). Though it measures college student mental health to a greater degree than most other comprehensive health surveys, feedback from survey participants, the literature, experts in the field, and college and university stakeholders suggests there are still many gaps in the mental health constructs measured. For example, though the literature indicates students are prescribed medication and undergo psychological treatment for a variety of mental health conditions (e.g., depression, anxiety, eating disorders) (Gallagher, 2007; Gilbert, 1992; Kitzrow, 2003), the original ACHA-NCHA instrument only assesses the degree to which these interventions are used for depression. Similarly, the original ACHA-NCHA does not ask about intentional self-harm behaviors, anger, and past-year diagnosis of a variety of prevalent mental health conditions.

Since December 2005, the ACHA-NCHA has undergone an extensive revisions process in which the scope of mental health issues measured has been dramatically revised and expanded. The intent of the revisions process to this national survey, which was grounded in theory and practices common to the field, was to create an instrument

that would expand upon the original version of the survey to elicit more comprehensive, timely, and accurate responses from today's college students in terms of their health needs and experiences. Because these data are used so extensively, because they are among the most commonly cited statistics associated with college student mental health, and because they are used for critical decision making purposes on campuses and nationally, it is imperative that the revised mental health indicators on the ACHA-NCHA be evaluated for their reliability and validity. As such, a rigorous analysis of the psychometrics of these revised indicators is both warranted and necessary.

Rationale for the Study

A review of the literature cites the great importance of survey data to understanding health and health behaviors in a variety of populations and for a variety of health and social issues (e.g., de Vaus, 1995; Fowler, 2002; Marsh, 1982; Di Iorio, 2005; Singleton & Straits, 1999; Trochim, 2006). Not surprisingly, surveys are among the most commonly utilized and applicable research tools used to best understand common health issues facing today's college students (Presley et al., 1994; Ouimet, Bunnage, Carini, et al., 2004). The development of a reliable and valid research instrument assessing a broad range of student health concerns (including mental health), which minimizes measurement error, is integral to the continued study of the field.

According to the ACHA-NCHA website (ACHA, 2007b), since 2000, data from the survey have been used by PSIs to:

- Identify the most common health and behavior risks affecting students' academic performance;

- Design evidence-based health promotion programs with targeted educational and environmental initiatives;

- Create social norms marketing campaigns by comparing students' actual behaviors to their perceptions about peer behavior;

- Allocate monetary and staffing resources based upon defined needs;

- Provide needs assessment data for campus and community task forces on sexual assault, alcohol use, eating disorders, etc;

- Produce graphs and data for policy discussions and presentations with faculty, staff, administration, and board members;

- Impact the campus culture by opening a dialogue about health with students and staff;

- Develop proposals to secure grant funding to expand or develop programs; and

- Evaluate programming efforts.

The ACHA (2007b) additionally claims that data are used by a variety of populations:

- By faculty for presentation in social sciences, health, communications, and research classes;

- By students for hands-on experience working with and analyzing data in courses, theses, and dissertations;

- By campus and local media for citations in articles and editorials;

- By administration for use in presentations for prospective students/parents and for freshmen orientation;

- By marketing professionals to draw on for promotional and health education materials.

Since the survey was first developed, more than 450,000 students on 400 unique campuses have participated in data collection efforts. Once only a paper-and-pencil survey, in Spring 2003, as a result of improved technological resources and demand by

the population, ACHA-NCHA administrators additionally began surveying students online. The number of PSIs who choose to administer the ACHA-NCHA each year continues to grow, as is evidenced in Table 1.1.

Perhaps there has never been a time in history in which the importance of a comprehensive, psychometrically sound measurement tool for student health—and one that measures mental health constructs in particular—has been warranted. With the current focus on evidence-based practice in public and mental health (Substance Abuse and Mental Health Services Administration [SAMHSA], 2007), in which approaches to prevention or treatment are validated by some form of documented scientific evidence, the need for tools with the ability to document this evidence (i.e., measure pre-/post-intervention differences) and monitor changes over time is necessary (SAMHSA, 2007). Further, as PSIs continue to concentrate on measurable student learning outcomes as well as effective assessment of whether those outcomes have been reached through PSI efforts (e.g., Suskie, 2006; American Association for Higher Education, 1991; Shavelson & Huang, 2003), the ACHA-NCHA has the capacity to serve as an invaluable tool. Lastly, with the national focus on prevention, detection, and intervention of mental health issues among college students (Voelker, 2007; NMHA & Jed, 2002; SPRC, 2004) the ACHA-NCHA, the only survey measuring mental health constructs as well as such a wide array of other risk behaviors among such a large sample of students and PSIs, is a unique and important source of data at both the institutional and national level. As such, continued growth of the survey effort is expected.

This study will contribute to the current public health research and evaluation knowledge base in a variety of ways. First, it will utilize current response analysis

techniques to begin to evaluate the degree to which mental health indicators on the revised ACHA-NCHA are reliable and valid. As described above, these data are used in a variety of unique and important capacities. Assessing the measurement quality of these indicators will provide important information to both the survey developers and the populations who utilize these data for their assessment and decision-making purposes. As these data are used for resource allocation, needs assessment, program development, and program evaluation, it is imperative they maximize truth and minimize error to every extent possible.

Second, the ACHA-NCHA measures a wide variety of college student health issues. As such, researchers commonly analyze these data to better understand the theoretical and practical relationships between a variety of health constructs (e.g., the relationship between high-risk drinking and depression, the relationship between demographic characteristics and suicide attempts). As mental health concerns on college campuses continue to rise, it is expected that researchers will continue to utilize these data for future analyses, publications, and scholarly activities. An assessment of the psychometric properties of these indicators will ensure measurement error for these survey items is minimized so the conclusions made in future research endeavors are as valid as possible.

Third, as mentioned throughout this introduction, the ACHA-NCHA provides a wealth of information on a variety of student health issues. This dissertation is examining the preliminary reliability and validity of only the revised mental health indicators. Hence, this research may serve as a template for future analyses of the reliability and validity of the other revised sections of the survey (i.e., health, health education, and

safety; alcohol, tobacco, and other drug use; sexual health; impediments to academic performance; and demographic characteristics), as well as for evaluating the reliability and validity of indicators on other large-scale surveys measuring public health constructs.

Lastly, though there are numerous published studies examining measurement and psychometric properties of both small- and large-scale health surveys and questionnaires, very few examine how changing the way researchers ask questions changes patterns in student responses (Presser et al., 2004), and experiments are needed to determine the effects of such changes. This study will provide an important statistical assessment of how instrumentation and measurement modifications impact the ways students respond to survey questions. Correspondingly, this research will add to the growing body of literature evaluating the best ways to collect reliable and valid self-report data on large-scale health surveys.

Research Questions and Hypotheses

Ultimately, as has been stated throughout this chapter, it was the intent of this researcher and of the entire ACHA-NCHA revisions committee to expand and modify the original ACHA-NCHA survey instrument to create a revised survey that better, more accurately, and more comprehensively captures today's students' health needs, behaviors, and experiences. Subsequently, it is anticipated that data collected from the revised instrument will demonstrate characteristics of being psychometrically sound. This study, which will address the reliability and validity of mental health indicators on a large scale survey measuring college student health, will evaluate the following research questions

and hypotheses using data collected during field pre-testing of the revised ACHA-
NCHA:

- **Research Question 1:** What patterns of student responses emerge on each of
  the revised ACHA-NCHA mental health indicators?

  Research Question 1A: How do student endorsements of each item vary
  according to demographic characteristics such as gender, year in school,
  and race/ethnicity?

- **Research Question 2:** What differences exist in student response patterns
  between comparable original ACHA-NCHA mental health items and revised
  ACHA-NCHA mental health items?

  Hypothesis 2: Changes in question formatting and response options
  between the original and revised ACHA-NCHA mental health indicators
  will result in significant differences in patterns of student responses on
  comparable items across versions of the survey.

- **Research Question 3:** To what extent are the mental health indicators on the
  revised ACHA-NCHA reliable?

  Hypothesis 3A: Measurement scales on the revised ACHA-NCHA will
  demonstrate sufficient levels of internal consistency in the pre-test sample.

  Hypothesis 3B: Reliability will not increase if any indicators are removed
  from measurement scales on the revised ACHA-NCHA in the pre-test
  sample.

- **Research Question 4:** To what extent are the mental health indicators on the
  revised ACHA-NCHA valid?

Hypothesis 4A: As a result of comparison with data collected from other surveys of mental health among college-aged students, the mental health indicators on the revised ACHA-NCHA will demonstrate evidence of criterion-related validity in the pre-test sample.

Hypothesis 4B: As a result of analyses demonstrating expected theoretical relationships between variables, mental health indicators on the revised ACHA-NCHA will demonstrate evidence of construct-related validity in the pre-test sample.

- **Research Question 5:** How do mental health indicators on the revised ACHA-NCHA differ from those on the original ACHA-NCHA in terms of their reliability and validity?

  Hypothesis 5: Mental health indicators on the revised ACHA-NCHA will demonstrate evidence of higher levels of reliability and validity than those on the original ACHA-NCHA.

Definition of Key Terms

The following is a list of terms used throughout this dissertation and their associated definitions.

- *Assessment*: Appraising the presence or magnitude of one or more personal characteristics. Assessing human behavior and mental processes includes such procedures as observations, interviews, rating scales, checklists, inventories, projective techniques, and tests (Aiken, 1997, p. 454).

- *College student***:** Any student enrolled (full- or part-time) in one of the more than 6,500 post-secondary institutions (PSIs) in the United States.

- *Construct validity***:** The evaluation of the validity of a measure by comparing results using that measure with the results expected on the basis of theory. If the results do not conform to theory, it is usually assumed that the measure rather than the theory is at fault (de Vaus, 1995, p. 389).

- *Criterion validity***:** The evaluation of validity by comparing results based on new measures of a concept with those using established measures (de Vaus, 1995, p. 389).

- *Field pre-testing***:** The evaluation of a survey instrument under realistic field conditions with respondents similar to those for whom the survey is designed (Singleton & Straits, 1999, p. 558).

- *Indicator***:** A specific measure of a more abstract concept (de Vaus, 1995, p. 390).

- *Internal consistency reliability***:** an 'equivalence' method of assessing reliability in which a statistical procedure is used to examine the consistency of scores across all the items constituting a measure (Singleton & Straits, 1999, p. 560).

- *Measurement***:** Procedures for determining the amount or quantity of some construct or entity (Aiken, 1997, p. 464).

- *Mental disorders***:** Health conditions that are characterized by alterations in thinking, mood, or behavior (or some combination thereof) associated with distress and/or impaired functioning (USDHHS, 1999).

- *Mental health***:** A state of successful performance of mental function, resulting in productive activities, fulfilling relationships with other people, and the ability to adapt to change and to cope with adversity. Mental health is indispensable to personal well-being, family and interpersonal relationships, and contribution to community or society (USDHHS, 1999).

- *Mental illness***:** A term that refers collectively to all diagnosable mental disorders (USDHHS, 1999).

- *Postsecondary education***:** Formal instructional programs with a curriculum designed primarily for students who are beyond the compulsory age for high school. This includes academic, vocational, and continuing professional education programs and excludes institutions that offer only avocational (leisure) and adult basic education programs." (Knapp et al., 2007).

- *Postsecondary institution(PSI)***:** An organization open to the public that has as its primary mission the provision of postsecondary education (Knapp et al., 2007).

- *Psychometrics***:** Theory and research pertaining to the measurement of psychological (cognitive and affective) characteristics (Aiken, 1997, p. 469).

- *Random measurement error***:** An error unrelated to the concept being measured that is the result of temporary, chance factors. Random errors are inconsistent across measurements (unpredictably varying in extent and direction) and affect reliability (Singleton & Straits, 1999, p. 565).

- *Reliability***:** The extent to which a psychological assessment device measures anything consistently. A reliable instrument is relatively free from errors of measurement (Aiken, 1997, p. 470).

- *Response analysis***:** A field pre-test technique in which the responses of pre-test respondents are tabulated and examined for problematic response patterns (Singleton & Straits, 1999, p. 566).

- *Response bias tendency***:** The tendency of a respondent to answer in a certain biased direction (such as in the direction of social desirability) as a function of the content or form of survey questions (Singleton & Straits, 1999, p. 566).

- *Scale***:** A composite measure of a concept constructed by combining separate indicators according to procedures designed to ensure unidimensionality or other desirable qualities (Singleton & Straits, 1999, p. 566).

- *Survey research***:** Any measurement procedures that involve asking questions of respondents (Trochim, 2006).

- *Systematic measurement error***:** Error from factors that systematically influence (bias) either the process of measurement or the concept being measured. Systematic errors are consistent across measurements taken at different times or are systematically related to characteristics of the cases being measured and thereby affect validity (Singleton & Straits, 1999, p. 569).

- *Validity***:** The extent to which an assessment instrument measures what it was designed to measure (Aiken, 1997, p. 474).

Summary

College students today face a variety of health risks, and research suggests the incidence and prevalence of mental health issues such as depression, anxiety, eating disorders, addiction, and suicide in this population have been increasing over the past few decades. The ACHA-NCHA, a comprehensive, self-report survey used to measure a variety of health constructs in this population, has served as one of the primary sources of data used to evaluate the current status of college students' health—and mental health in particular—since 2000. As a result of feedback from a variety of sources and current survey development research and theory, survey administrators, including this researcher, began the process of revising the instrument in December 2005. This revisions process resulted in dramatic changes and additions to the items and indicators measuring student mental health, each of which was intended to create the most psychometrically sound survey instrument possible. The revised ACHA-NCHA was field pre-tested with students from seven PSIs in Spring 2007. This dissertation is a preliminary response analysis and assessment of the reliability and validity of these revised mental health indicators.

CHAPTER II: LITERATURE REVIEW

Introduction

This chapter presents a review of the literature relevant to: (1) the extent and significance of college student mental health concerns, including the determinants of mental health problems in this population; (2) an overview of the applicability and use of surveys, and particularly online surveys (i.e., the method used by the ACHA-NCHA field pre-test), to assess college health; (3) the Model of Survey Response, which provides the theoretical foundation for examining how people respond to questions they are asked on surveys and lays the foundation for developing reliable and valid survey questions; (4) characteristics of reliable and valid survey questions, which guided many of the ACHA-NCHA revisions and addtions; (5) issues and practices in establishing reliability and validity in survey research; and (6) issues and practices in pre-testing survey questionnaires.

Extent and Significance of College Student Mental Health Concerns

The current literature on college students' mental health suggests there has been an increase in the prevalence of mental health disturbances and distress in this population over the past few decades (e.g., Sharkin, 2006; Gallagher, 2007; Gilbert, 1992; Kadison & DiGeronimo, 2004; Kadison, 2006). Many PSI administrators and health professionals today claim they see fewer cases of students who present with traditional developmental struggles and more who present with serious psychological problems (Sharkin, 2006). For example, Benton and colleagues (2003), who conducted a 13-year study from 1989-

2001 of 13,257 students seeking help at a large Midwestern university counseling center, found that:

> "students [today] experience more stress, more anxiety, and more depression than a decade ago and … [s]ome of these increases were dramatic. The number of students seen each year with depression doubled, while the number of suicidal students tripled, and the number of students seen after a sexual assault quadrupled" (p. 69).

The 2007 National Survey of Counseling Center Directors, which surveyed 272 counseling directors from colleges and universities across the country, found that 91.5 percent of directors report this recent trend toward greater numbers of students with severe psychological problems appears to be true on their campuses (Gallagher, 2007). Further, 60 percent of university deans reported that students are using psychological services at record rates and for longer periods of time than in years past (Levine & Cureton, 1998a).

Psychologists, counselors, and mental health researchers have cited countless causes for this increase in pathology. At the individual level, it has been suggested that because of the increasing competition to gain admission to college, students are entering the environment more overwhelmed and more "damaged" than in years past (Shea, 2002; Levine & Cureton, 1998b). At the familial level, students today are more likely to have experienced divorce, family dysfunction, poor parenting skills, and instability than previously, all of which are risk factors for mental illness (e.g., Kadison & DiGeronimo, 2004; Gallagher, 2007). At the college- and university-wide level, demographics of students attending PSIs have also changed. In the 1980s, the number of U.S. high school students declined, and colleges and universities began recruiting nontraditional students, focusing on graduate, older, and international students (Brindis & Reyes, 1997). These students experience unique stressors, are more likely to experience a mental health

disorder, and require additional mental health support than traditional students (Brindis & Reyes). Additionally, the number of people who are seeking a college education is increasing, and some of those who are now seeking higher education are poorly prepared for the academic demands and adjustment required to be successful in college, which may contribute to the development of a variety of mental health issues (Cook, 2007). At the societal level, efforts to decrease the stigma associated with mental illness have likely been effective, and more students are requesting services (Padron, 2006; Gilbert, 1992) and are being referred to services by faculty and staff (SPRC, 2004) than ever before.

Concurrently, there has also been a significant movement in the mental health field to better understand the biological basis of mental health disorders in the last 20 years (Kadison, 2006). This has allowed for more successful assessment, intervention, and management of psychiatrically ill adolescents during high school, allowing them to further their educations and enroll in PSIs (Kadison, 2006; SPRC, 2004). Specifically, in the past two decades, a new generation of safe, effective medications have been developed that have a wide range of therapeutic benefits to ameliorate depression, anxiety, panic, social anxiety, eating disorders, and obsessive compulsive disorder (e.g., Kadison, 2006; Caulfield, 2001; Geddes, Freemantle, Harrison, & Bebbington, 2000). Recent estimates suggest there has been an 800 percent increase in the prescribing of antidepressants during the decade following the 1988 release of Prozac (Caulfield, 2001), and 87.5 percent of PSI counselors believes there has been a significant increase in the percentage of students who come to campus on medication in recent years (Gallagher, 2007). Counseling center directors report 23.3 percent of their clients today are on psychiatric medication, which is a marked increase from 20 percent in 2003, 17 percent

in 2000, and 9 percent in 1994 (Gallagher). As such, more students are entering the college environment today with a history of mental illness diagnosis and treatment (Mowbray et al., 2006), and this history of mental health disorders is a significant risk factor for distress while in college (Shea, 2002).

Although the psychological problems of college students today are believed to be more serious and more prevalent than in decades past, the literature suggests there have always been numerous aspects of the college environment and experience that contribute to the development of mental health disorders in this population (Sharkin, 2006). First, the age at which most people experience symptoms of mental disorders, usually cited to be between ages 16-25 (e.g., NMHA & Jed, 2002; Collins & Mowbray, 2005; Kessler et al., 1995; Kessler, Chiu, & Walters, 2005), correlates almost perfectly with the age at which students typically seek higher education. Second, college is a time during which students typically develop their racial, religious, and sexual identities, which often present developmental and maturational challenges (Kadison & DiGeronimo, 2004; Sharkin, 2006). Not surprisingly, then, the stress and turmoil that accompany such major developmental milestones is a factor in the development of mental health disturbances in these students (NMHA & Jed, 2002). Kadison & DiGeronimo state:

> "When kids go off to college, society expects that their identity will shift from being dependent children to being responsible adults, but,…this expectation and the reality of the experience often clash—and for a good reason. The societal pressure to become more autonomous and independent comes at a time when college students are entering a new world where they need extra help and guidance…the reality is that they don't yet have other sources of comfort and guidance, and so the stress of being 'grown up' is magnified" (pp. 12-13).

Lastly, it has been suggested that he dynamic relationship between the person and environment in stress perception and reaction is especially magnified in college students (Ross et al., 1999), largely because in addition to developmental stressors, there are also a

number of stressors unique to the collegiate environment. These include examinations, public speaking, interpersonal relationships, sharing a confined space with new people, meeting the expectations or demands of parents, coping with family problems, handling long-distance relationships with significant others, balancing responsibilities, financial problems, peer pressure, and ultimately graduating and beginning one's career (e.g., Grace, 1997; NMHA & Jed, 2002; Oswalt & Finkelberg, 1995). Table 2.1 provides an overview of the common causes of stress reported in a recent study of 100 undergraduates at a Midwestern university, by type of stressor—interpersonal, intrapersonal, academic, and environmental (Ross et al., 1999). Developing healthy and effective coping mechanisms to manage these stressors is imperative for the preservation of a student's mental health. When students are unable to cope with these unique stressors, or do so ineffectively, they are at risk for developing a variety of mental health disturbances (Cook, 2007).

All of the factors described above, including increasing prevalence rates and severity of mental illness on college campuses, an increased awareness of the biological mechanisms underlying mental illness, developmental challenges associated with the college experience, and stressors reported by students today and throughout history, have led to an ever-expanding understanding and comprehensive body of literature examining college student mental health. Although students may be susceptible to any mental health challenge or diagnosis experienced in the human condition, the literature suggests they are particularly at risk for the following issues and/or diagnoses, each of which is briefly described below: anxiety, depression, suicidal behavior, deliberate self-inflicted harm, substance abuse, and eating disorders. Though each is discussed independently, it is

Table 2.1. Stressors experienced by college students. Adapted from Ross et al. (1999).

Type of stressor	Percent of students reporting stressor*
Academic	
• Increased class workload	73
• Lower grade than anticipated	68
• Change of major	24
• Search for graduate school or job	21
• Missed classes	21
• Anticipation of graduation	20
Intrapersonal	
• Change in sleeping habits	89
• Change in eating habits	74
• New responsibilities	73
• Financial difficulties	71
• Held a job	65
• Spoke in public	60
• Change in use of alcohol or drugs	39
• Outstanding personal achievement	35
• Started college	32
• Decline in personal health	26
Interpersonal	
• Change in social activities	71
• Roommate conflict	61
• Work with people you don't know	57
• Fight with boyfriend/girlfriend	41
• New boyfriend/girlfriend	36
• Trouble with parents	21
Environmental	
• Vacations/breaks	82
• Waited in long line	69
• Computer problems	69
• Placed in unfamiliar situation	51
• Messy living conditions	50
• Put on hold for extended period of time	47
• Change in living environment	46
• Car trouble	42

*n=100 students

Note: Only those stressors reported by 20 or more students are reported in this table; additional stressors were present in this sample. See Ross et al., (1999) for additional information.

important to remember they often co-occur in this population, have many of the same underlying determinants, and often require similar types of intervention and treatment (Kadison & DiGeronimo, 2004; Sharkin, 2006).

Anxiety

As they affect more than 40 million Americans aged 18 and older (NIMH, 2007b) and have a lifetime prevalence of 30.5 percent (Kessler et al., 20005), anxiety disorders are the most commonly experienced mental illnesses in the United States. Anxiety, a feeling which is generally characterized as diffuse and unpleasant with a sense of apprehension or worry (Bhave & Nagpal, 2005), is a very normal and appropriate reaction to many life stressors and is certainly to be expected in an academically rigorous and challenging environment such as college. However, anxiety can become a disorder when symptoms are severe, pervasive, and persistent; when they interfere with normal life (Bhave & Nagpal); and/or when they last longer than 6 months (NIMH, 2007b).

Anxiety disorders are most commonly manifested as generalized anxiety disorder (GAD), obsessive-compulsive disorder (OCD), panic disorder, posttraumatic stress disorder (PTSD) and phobias (Anxiety Disorders Association of America [ADAA], 2007). They are commonly associated with physical symptoms such as headaches, muscle tension, perspiration, restlessness, and chest and stomach discomfort (Bhave & Nagpal, 2005), as well as additional symptoms of excessive and often irrational worry, tension and irritability, fear, difficulty concentrating and making decisions, and increased avoidance and social withdrawal (NIMH, 2007b).

Anxiety disorders, which are more common in women than in men, commonly develop in adolescence, and 75 percent of all people with an anxiety disorder will

experience symptoms before they are 22 years old (NIMH, 2007b). Accordingly, it is not surprising that many students first experience signs of maladaptive anxiety while enrolled in college. In 2000, it was estimated that nearly 7 percent of American college students experienced symptoms of an anxiety disorder within the past year, and panic disorder is frequently cited as a top reason for women dropping out of college (ADAA, 2007). The surveillance, prevention, diagnosis, and treatment of anxiety disorders in college students is necessary on college campuses, particularly among college women.

Depression

Depression is typically associated with a feeling of sadness and gloom, often with reduced activity (Bhave & Nagpal, 2005). Feeling sad or depressed is a normal reaction to a tragedy, change, or significant loss, and for most people, these symptoms are temporary. However, a depressive disorder may be present when these feelings are experienced for a prolonged period of time and/or and accompanied by diminished motivation, low self-esteem, low energy, and impaired thinking and emotional well-being (Bhave & Nagpal).

Depressive disorders are part of a larger classification of mood disorders, which affect nearly one in five Americans in their lifetimes (Kessler et al., 2005). Depression is more common among women than among men, and biological, life cycle, hormonal and psychosocial factors unique to women may contribute to this higher rate (NIMH, 2007c). Depression has also been shown to be more prevalent in the college population than in the general population (Leino & Kisch, 2005), and evidence suggests more students than ever before have been diagnosed with depression. Estimates from the Spring 2005 ACHA-NCHA (ACHA, 2006b) suggest that approximately 19.2 percent of females and

10.8 percent of males have been diagnosed with depression in their lifetimes, which is a marked increase from the 12.8 percent of females and 6.2 percent of males who reported ever being diagnosed with depression in Spring 2000 (Leino & Kisch, 2005). Unfortunately, many students who are diagnosed with depression do not receive appropriate treatment or any treatment at all (Kadison & DiGeronimo, 2004). Additionally, the National Institute of Mental Health (2007c) reports that depression in adolescence frequently co-occurs with other disorders such as anxiety, disruptive behavior, eating disorders, or substance abuse and can contribute to increased risk for suicide.

Suicidal behavior

Over the past decades, the rates of suicide have been increasing dramatically among young people (Lake & Tribbensee, 2002). Suicide is currently the second leading cause of death in U.S. college students, surpassed only by accidents (SPRC, 2004). Results of the Big 10 Suicide Study indicate suicide occurs at a rate of 7.5 per 100,000 students (Silverman, Meyer, Sloane, Raffel, & Pratt, 1997), which equates to nearly 1,100 student deaths from suicide each year (NMHA & Jed, 2002). The largest number of suicides for both males and females was in the 20–24 year age group (46 percent) and among graduate students (32 percent) (Silverman et al.). Table 2.2 depicts rates of suicide in college students by gender and age group as provided by Appelbaum (2006).

Table 2.2. Suicide rates of college students (per 100,000). Adapted from Appelbaum (2006).

Group	Women	Men	Total
Underclassmen (ages 17-19)	1.2	5.7	3.4
Upperclassmen (ages 20-24)	4.5	9.0	7.1
All students (ages 17 to 49)	4.5	10.0	7.5

Suicide has been described as the end of a continuum that begins with suicidal ideation, continues with planning and preparing for suicide, and ends with threatening, attempting, and completing suicide (SPRC, 2004). When these behaviors are taken into account, the problem of suicide becomes even larger in the college population. Nearly one in ten college students reported seriously considering suicide in the last school year (ACHA, 2006b), one in twelve report making a suicide plan (NMHA & Jed, 2002), and 1.5 percent reported an actual suicide attempt (ACHA, 2006b).

A current culture of ambition, high anxiety, and increased rates of depression are thought to contribute to high lifetime prevalence of suicide ideation among college students (Sontag, 2002; Meilman & Pattis, 1994). One recent study found 90 percent of suicide attempters had experienced a work or school failure, and 46 percent of those who threatened suicide had relationship difficulties or a recent breakup (Meilman & Pattis, 1994). Suicidal behaviors are also correlated with alcohol use, and as many as 27 percent of suicide attempts and 21 percent of suicide threats may involve alcohol consumption (Meilman & Pattis). Therefore, addressing these issues through preventive programming may be important to help reduce suicide among students.

Deliberate self-harm

Deliberate self-harm (DSH)—the deliberate, direct destruction or alteration of body tissue without conscious suicidal intent, but resulting in injury severe enough for tissue damage to occur (Gratz, Conrad, & Roemer, 2002)—is a relatively new area of study in the public health literature. Though there is only a small body of knowledge available regarding the prevalence, nature, and correlates of DSH in a college population,

there is a widely held belief in the field that the prevalence of these behaviors has increased in recent years (Whitlock, Eckenrode, & Silverman, 2006; Sharkin, 2006).

In the largest DSH study of college students to date (n=2,875), Whitlock and colleagues (2006) found the lifetime prevalence of DSH in college students was 17.0 percent, and the 12-month prevalence of DSH was 7.3 percent. Of those who had engaged in self-harm behaviors, nearly three in four students had done so repeatedly. Further, 21.1 percent of those who had practiced DSH had injured themselves more severely than they intended (Whitlock et al.). The specific DSH behaviors most frequently reported in the literature among this population are needle sticking, skin cutting, scratching and head banging (Whitlock et al; Gratz et al., 2002; Brown, Williams, & Collins, 2007).

Brown and colleagues (2007) suggest that self-harm behaviors are often utilized when individuals feel they have no other way to express or control negative emotions. For these students, DSH may serve to assist them in expressing negative emotions, decrease negative affect, and/or avoid unpleasant emotions (Brown et al.). There are several risk factors hypothesized for DSH, including dissociation, sexual abuse, childhood separation, emotional neglect, physical abuse, and insecure attachment (Gratz et al., 2002). Current literature suggests no gender differences (Brown et al; Gratz et al.) or only minimal gender differences (Whitlock et al., 2006) in risk for DSH in college populations.

Deliberate self-harm has been associated with increase risk for distress, suicidal ideation and attempts (Whitlock et al., 2006), and poor management of negative emotions and avoidance (Gratz et al., 2002) in college students. Perhaps most alarming about this

college health concern is the low frequency at which students disclose these behaviors:

recent estimates suggest 40 percent of self-harmers kept their behavior entirely to

themselves and only one-quarter of students discussed their DSH behaviors with a mental

health provider (Whitlock et al., 2006). Not surprisingly, then, the literature suggests

there is a critical need for surveillance, detection, intervention, and prevention of these

behaviors (Gratz et al., 2002; Whitlock et al.; Brown et al., 2007).

Substance abuse

Perhaps no health issue among college students has been researched and

documented in the literature as extensively as substance, or alcohol and other drug, abuse.

Underage and excessive drinking, as well as other substance abuse, by college students

has long been recognized as a major problem in higher education (U.S. Department of

Education, 2007). It has been suggested that approximately 80 percent of college students

drink alcohol, that college students engage in high-risk alcohol use with more frequency

than their non-college peers, and that high-risk use appears to be climbing (National

Institute of Alcoholism and Alcohol Abuse [NIAAA], 2002). A recent national study

found the percentage of students at four-year institutions who said they drink "to get

drunk" rose from 40 percent in 1993 to 48 percent in 2001 (Weschler, Lee, Kuo, et al.,

2002). Nearly 600,000 students between the ages of 18 and 24 are unintentionally injured

each year under the influence of alcohol, and approximately 1,700 college students die

each year from these alcohol-related unintentional injuries, including motor vehicle

crashes (Hingson, Heeren, Winter, & Wechsler, 2005). High-risk alcohol use has been

associated with a variety of consequences in college students, including but not limited

to: assault, sexual abuse, unsafe sex, academic difficulties, drunk driving, vandalism, and

police involvement (e.g., Hingson & Howland, 2002; Hingson et al., 2005; Weschler et al., 2002; Presley et al., 1998).

Alcohol and other drug abuse in college students is linked to very serious health concerns in this population. More than 150,000 students develop an alcohol-related health problem each year (Hingson et al., 2002) and studies suggest that between 1.2 and 1.5 percent of students indicate that they tried to commit suicide within the past year due to drinking or drug use (Presley et al., 1998). McCarty and colleagues (2004) found that 50 percent of men and 44 percent of women who were binge drinkers at 17-20 years continued to be binge drinkers at 30-31 years of age; therefore, problem drinking in adolescence predicts problem drinking in adulthood. The literature also suggests that 31 percent of college students meet criteria for a diagnosis of alcohol abuse and 6 percent for a diagnosis of alcohol dependence (i.e., physiological dependence) in the past 12 months (Knight, Weschler, Kuo, et al., 2002). Men are more likely to report high-risk drinking on college campuses and are also more likely to meet criteria for abuse and dependence (Presley et al., 1998)

Illicit substance use is also of concern among college students in the United States. Results from the Monitoring the Future Study (Johnson, O'Malley, Bachmnan, & Schulenberg, 2006) found that the percentage of college students who reported any illicit substance use in the last year rose from 29.2 percent in 1991 to 36.6 percent in 2005, and 19.5 percent of students used illicit substances in the past 30 days. Results also indicate that, in 2005, 33.3 percent reported using marijuana, 5.0 percent reported using hallucinogens, 2.9 percent reported using ecstasy, 5.7 percent reported using cocaine, and 0.3 percent reported using heroin. Survey results additionally show that 8.2 percent of

current college students report using other narcotics such as Vicodin, OxyContin, and Percocet (Johnson et al., 2006).

As many as 12 percent of college students believe they have a substance abuse problem (Presley et al., 1998). As substance abuse is very likely to co-occur with other mental health disorders (Kessler et al., 2005), particularly among men, it is critical that PSIs continue to monitor rates of alcohol and other drug use on their campuses as well as develop effective prevention and intervention strategies.

Eating disorders

People with eating disorders experience serious disturbances in their eating patterns, such as a severe and unhealthy reduction in their food intake or overeating, as well as extreme concern about body shape or weight (NIMH, 2007d). The most common forms of eating disorders in the United States are anorexia nervosa (AN), bulimia nervosa (BN), and binge eating disorder (BED). AN is characterized by emaciation, a relentless pursuit of thinness and unwillingness to maintain a normal or healthy weight, a distortion of body image and intense fear of gaining weight, a lack of menstruation among girls and women, and extremely disturbed eating behavior. BN is characterized by recurrent and frequent episodes of eating unusually large amounts of food (e.g., binge-eating), and feeling a lack of control over the eating. This is followed by behavior that compensates for the binge, such as purging (e.g., vomiting, excessive use of laxatives or diuretics), fasting and/or excessive exercise. Unlike AN, people with BN usually fall within the normal range for their age and weight; however, they often fear gaining weight, want desperately to lose weight, and are intensely unhappy with their body size and shape. Usually, bulimic behavior is done secretly, because it is often accompanied by feelings of

disgust or shame. The binging and purging cycle usually repeats several times a week. Lastly, BED is characterized by recurrent binge-eating episodes during which a person feels a loss of control over his or her eating. Unlike BN, binge-eating episodes are not followed by purging, excessive exercise or fasting. As a result, people with binge-eating disorder are often overweight or obese. They also experience guilt, shame and/or distress about the binge eating, which can lead to more binge-eating (NIMH, 2007d).

In general, eating disorders are far more common in females than males, though studies in the last ten years suggest the prevalence of eating disorders and body image distortion in males is significantly increasing (e.g., Braun, Sunday, Huang, & Halmi, 1999). Men and boys account for only an estimated 5 to 15 percent of patients with anorexia or bulimia and an estimated 35 percent of those with binge-eating disorder (NIMH, 2007d). In a review of incidence and prevalence rates of eating disorders in the United States, Hoek and van Hoeken (2003) found that the average prevalence rate for anorexia nervosa was 0.3 percent for young females and the prevalence rates for bulimia nervosa were 1 percent and 0.1 percent for young women and young men, respectively. The estimated prevalence of binge eating disorder is approximately 1 percent. The National Eating Disorders Association (2005) suggests this means as many as 10 million females and 1 million males are currently struggling with a diagnosis of AN or BN, and 25 million more are struggling with binge eating disorder.

College students have been well documented as a high-risk group for eating disorders (e.g., Phillips & Pratt, 2005; Schwitzer, Bergholz, Dore, & Salimi, 1998; Woodside & Garfinkel, 1992; Drewnowski, Yee, Kurth, & Krahn, 1994), which are often related to constructs such as ineffectiveness, asceticism, poor impulse regulation,

perfectionism, poor self esteem, interpersonal distrust, social insecurity, and maturity fears (Phillips & Pratt, 2005). Various biological, cultural, and environmental factors including messages from the media, unrealistic beauty standards, ineffective coping, desire for acceptance, and need for control impact the likelihood that one will develop an eating disorder while in college. When women feel a lack of control in their new college environments, they are often susceptible to eating disorders and depression (Beeber, 1999).

Studies suggest that about one third of all women who develop eating disorders do so in college (Woodside & Garfinkel, 1992), and as many as 25 percent of college women are thought to be at risk of developing eating disorders (Drewnowski et al., 1994). Further, as many as 91 percent college women have attempted to control their weight through dieting and 22 percent dieted "often" or "always" (Kurth, Krahn, Nairn, & Drewnowski, 1995). More than one third (35 percent) of these "normal dieters" progress to pathological dieting, and of those, 20-25 percent progress to partial or full syndrome eating disorders (Shisslak, Crago, & Estes, 1995).

For females aged 15-24, the mortality rate of AN is 12 times higher than the death rate of all other causes of death and AN has the highest premature fatality rate of any mental illness (Sullivan, 1995). Unfortunately, only a minority of the people who meet stringent diagnostic criteria for eating disorders receive mental health care, which suggests the majority of persons with a severe eating disorder lack adequate treatment (NIMH, 2007d). As such, colleges and universities have been charged with the task of effective monitoring, prevention, and intervention with these mental disorders on their campuses, particularly among their female students.

Summary

A variety of factors contribute to the development of mental health disorders in college students today including developmental and maturational challenges, changing student demographics, stress, and environmental aspects of the college environment. Students are particularly likely to experience anxiety, depression, suicidal ideation and attempts, deliberate self-harm, substance abuse, and eating disorders. Recent events such as high-profile student suicides, the Virginia Tech massacre, and the Northern Illinois University tragedy have highlighted the need to monitor, address, treat, and prevent mental health crises among students and the devastating consequences that can result on campuses nationwide. As higher education remains a national priority, it is imperative that students' health be preserved and protected. As such, further research into student mental health determinants, needs, and experiences is warranted and needed. Surveys provide one mechanism through which this research may be conducted.

The Applicability and Use of Surveys to Assess College Health

Perhaps no other research tool is used as extensively as the survey in behavioral and social research (Bowden, Fox-Rusby, Nyandieka, & Wanjau, 2002). Surveys, or special types of questionnaires designed to address specific research questions, are one of the primary methods used in social sciences to test hypotheses and discover relationships between theoretical constructs (Marsh, 1982). Many historians have traced the beginning of social survey research to the end of the 19th and early 20th centuries with the most dramatic growth in the field occurring in the 1960s and 1970s, largely due to the rapid expansion of telephone-assisted surveys (Marsh). Though the federal government is the

largest collector of survey data in the U.S. (Fowler, 2002), thousands of surveys are

conducted each year by university, non-profit, and commercial entities.

The basic characteristics of surveys, as outlined by Fowler (2002, pp. 1-2), are as

follows:

- The purpose of the survey is to produce statistics, that is, quantitative or numerical descriptions about some aspect of the study population;
- The primary way of collecting information is by asking people questions, and their answers constitute the data to be analyzed;
- Generally, information is collected about only a fraction of the population (i.e., a sample), rather than from every member of the population.

Correspondingly, the primary elements of a survey questionnaire are the questions

themselves, the response formats or categories that accompany the questions, and any

special instructions that appear in the questionnaire or that are associated with a particular

question to tell the respondent how to address it (Fowler, 2002).

Because surveys are used so extensively in the social sciences, there is a

tremendous body of literature, including numerous textbooks (e.g., Fowler, 2002; Aday

& Cornelius, 2006), on principles of their development, construction, validation, and

implementation. The very basic steps in survey development are (1) state the general

purpose of the survey, (2) write the objectives the survey is attempting to accomplish, (3)

select and write items that will be used to gather data to accomplish objectives, (4)

determine overall survey format, and (5) write rules for scoring and administration (Di

Iorio, 2005).

A goal of numerous surveys is to better understand a variety of health phenomena,

and current literature provides an extensive overview of the applicability of using survey

methods to measure health constructs in a variety of populations. Aday and Cornelius

(2006) explain,

"Health surveys serve as a critical resource to measure the health status, risk factors, and health behaviors of the population and to assess the level of quality of the health care received. They also permit the identification of disparities in health care associated with access, use, cost, and insurance coverage and serve to identify related patterns and trends over time. The descriptive and analytical findings they generate are key inputs to facilitate the development, implementation, and evaluation of policies and practices addressing health and health care" (p. xi).

Researchers can collect health survey data through a variety of channels, including face-to-face interviews, telephone, paper-and-pencil surveys, and computer-assisted administration. Regardless of survey method, self-report is the most common form of measure used to collect data about health and health behaviors (Di Iorio, 2005). Though this method of data collection is prone to errors such as recall and social desirability biases, which are described in more detail below, self-report is essential as respondents are often the only individuals who know their personal experiences in regards to a particular issue (Di Iorio). Further, even when other data on phenomena exist, they may be too difficult or too costly to obtain (Baldwin, 2000).

In addition, health behavior and health education often depend on the measurement of abstract concepts and constructs, such as attitudes and beliefs, which can typically only be measured through self-report (Di Iorio, 2005). In these cases, measurement must be extended to include the process of converting underlying, latent concepts into empirical indicators. Surveys provide an opportunity for this operationalization in the use of measurement scales. Scales are defined as collections of items combined into a composite score intended to reveal levels of theoretical variables not readily observable by direct means (DeVellis, 2003). Though not every concept requires a multi-item scale on a health survey, the ability to combine single-indicator

questions and multiple-indicator scales on a single instrument to measure a variety of constructs is one of the greatest strengths of survey methodology.

Recent technological advances, including the growth of the World Wide Web (WWW), have expanded the use of surveys in a variety of capacities. The use of the Internet as research tool emerged during the mid-1990s (Epstein & Klinkenberg, 2001), during a transitional period in survey research in which investigators were experiencing increasing costs and declining response rates (Couper, 2007). The potential of the WWW to offer access to vast numbers of subjects—including those with rare characteristics—as well to administer surveys rapidly, with reduced cost, and with designs bolstered by computer logic intrigued the communication scientists, sociologists, and psychologists who pioneered these e-survey methodologies (Couper; Eysenbach & Wyatt, 2002).

The use of web-based surveys has offered a unique and meritorious contribution to the health field in particular (e.g., Eysenbach & Wyatt, 2002; Pealer et al., 2001; Pealer & Weiler, 2000; Karras & Tufano, 2006; Daley, McDermott, McCormack Brown, & Kittleson, 2003; Whitehead, 2007), and it has been suggested that data gathered from web-based surveys have the potential to improve both the effectiveness and efficiency of health intervention design, deployment, and evaluation (Karras & Tufano, 2006). Research indicates that reliability levels are comparable between online and paper instruments for a variety of health constructs (e.g., Herrero & Meneses, 2006), and that respondents may report sensitive behaviors and events (e.g., sexual behaviors, high-risk alcohol use) with less bias on web-based surveys than on traditional self-administered pen-and-paper surveys or interviews (e.g., Di Iorio, 2005; Aday & Cornelius, 2006).

In a meta-analysis of 61 studies comparing computer to pencil-and-paper versions of the same instrument, researchers found that social desirability distortion bias was lower when using computer-based methodologies, especially when respondents were alone and could respond freely (Richman, Kiesler, Weisb, & Drasgow, 1999). Herrero and Meneses (2006) summarize key reasons for this phenomena, including (1) the absence of an experimenter engaged in face-to-face interaction with the participant reduces bias for a correct assessment of highly sensitive psychological issues, and (2) participants immersed in interaction with a computer-based assessment may feel a sense of invulnerability to criticism, an illusion of privacy, and an impression that their responses disappear into the computer once the assessment is complete. These advantages are similar to others documented in the literature, which include speed, asynchronous communication, absence of intermediaries, and ephemerality (Thatch, 1995, as cited in Daley et al., 2003).

Despite the numerous advantages of web-based surveying, there are also limitations. Technological difficulties such as variations in hardware configurations, browsers, operating systems, and end-user settings (Karras & Tufano, 2006) often make the process of e-survey development quite cumbersome. Further, though comparable to traditionally mailed surveys, web-based surveys frequently experience moderately low response rates, which may limit generalizability of results (Couper, 2007). Additionally, with internet-based surveys, data collection environments can neither be controlled nor monitored (Daley et al., 2003), which has the potential to introduce unique sources of error into survey measurement.

Though there are a variety of challenges associated with it use and it is still very much in its infancy (Daley et al., 2003), e-surveying is expected to experience continued expansion in the health field in coming years (e.g., Eysenbach & Wyatt, 2002; Whitehead, 2007). As the literature on this method continues to evolve, researchers are providing more evidence-based guidance on how to best design web-based surveys. Some very basic principles of web survey design are provided below in Figure 2.1.

Fundamental Principles of Web Survey Design

- Provide specific instructions on how to execute the technological action that is needed to flow through the survey
- Use question formats that have known measurement problems in paper questions (e.g., select all that apply) sparingly
- Present each question in a way that is similar to what is typically used in a paper self-administered survey
- Limit the use of color to encourage consistency in appearance and improve the navigational flow through the web survey
- Use drop-down boxes sparingly
- Do not require (i.e., force) a respondent to answer a question before going to another one
- Develop skip patterns that allow respondents to click on a link that takes them directly to the next appropriate question

---

Figure 2.1. Principles of web-based survey design. Derived from Dillman (2000).

The use of surveys, and particularly online surveys, is especially relevant in PSI settings. Surveys are the most commonly utilized research tool to gauge and monitor student experiences on college campuses, and student self-reports about the nature and frequency of their behaviors collected through surveys have been considered accurate indicators of these activities (Ouimet Bunnage, Carini, Kuh, & Kennedy, 2004). Additionally, college and university students have higher levels of internet access than the general population (Jones & Madden, 2002) and are generally thought to be more

technologically-savvy, which may make the use of e-surveys especially applicable. It has also been suggested that students find web-based surveys less cumbersome, quicker, and more convenient to their hectic lives than paper-and-pencil questionnaires (e.g., Ouimet et al., 2004).

Because of their relative ease of administration and ability to meet many of the needs of their respondents, as well as their general utility to measure a variety of sensitive health constructs, there is a reasonable body of literature citing the applicability of online surveys to monitor and measure college students' health (e.g., Daley et al., 2003; Pealer et al., 2001; Kypri, Gallagher, & Cashell-Smith, 2004). Current literature suggests, however, that most of these surveys are localized to individual campuses or consortia and have not been implemented nationwide or on a large scale (Ouimet et al., 2004).

The Core Alcohol and Drug Survey (Core Institute, 2007) and the ACHA-NCHA are currently the only two large-scale national online surveys measuring student health constructs found in the literature. Hence, the ACHA-NCHA is the only large-scale online survey measuring a wide array of health constructs, including health, health education and safety; alcohol, tobacco, and other drug use; sexual health; mental and physical health; impediments to academic performance; and demographic characteristics. Yet, gaps in the mental health section of the ACHA have been noted by institutions that utilize the survey as well as the literature. Soet and Sevig (2006) state:

> "…the ACHA has conducted the National College Health Assessment (NCHA), an informative survey of college student health that includes limited questions on mental health issues such as medication use, depression, and suicide. Although these efforts have provided us with some information about mental health issues facing college students, little is known about the breadth and depth of mental health issues on college campuses…[I]nformation is needed about the broad range of mental health challenges within the context of the college student life beyond depression and suicide…" (p. 411).

Because this feedback emerged as a theme from various sources and because of the emergent trends in college mental health described earlier in this chapter, the ACHA-NCHA revisions committee placed special emphasis on expanding this section of the survey during its revisions process. As the intent of this dissertation is to evaluate the reliability and validity of the revised and newly introduced items that constitute the expanded mental health section, the remainder of this literature review concentrates on the Model of Survey Response, which provides theoretical underpinnings to understand how people answer questions on surveys; general characteristics of reliable and valid survey questions; issues and practices in establishing reliability and validity in survey research; and a review of current strategies used to pre-test and assess reliability and validity in revised survey questionnaires.

Model of Survey Response

As the field of survey development expanded immensely toward the last quarter of the 20th century, a great deal of research emerged to better understand how people answered the questions asked on these instruments. These endeavors, which spawned the development of survey response theory, created a great paradigm shift in the field—that from the "art" of survey development to the "science" of survey development—upon which research continues to expand today (Fowler, 1995; Fowler, 2004). With roots and perspectives in cognitive psychology, these emergent theories provide guidance on how to design and evaluate questions and questionnaires so as to reduce error and maximize truth in survey efforts (Bowden et al., 2002; Aday & Cornelius, 2006; Fowler, 2004).

Perhaps the most commonly cited and widely accepted theory that emerged during this time was Tourangeau's (1984) Model of Survey Response. This framework suggests the response process a person uses when answering survey questions is divided into four primary stages—comprehension of the item, retrieval of relevant information, use of that information to make required judgments, and selection and reporting of an answer (Tourangeau, 1984; Tourangeau, Rips, & Rasinski, 2000). These components and the specific cognitive processes used in each are described in Table 2.3.

Table 2.3. Components of the response process. Derived from Tourangeau and colleagues (2000, p. 8).

Component	Specific processes used to answer survey questions
Comprehension	• Think about questions and instructions • Grasp logical form of the question • Identify question focus and information sought • Link key terms to relevant concepts
Retrieval	• Generate retrieval strategy and cues for retrieval • Retrieve specific, generic memories • Fill in missing details
Judgment	• Assess completeness and relevance of memories • Make inferences based on accessibility • Integrate material retrieved • Form estimate based on partial/incomplete record
Response	• Map judgment onto response category • Modify response as necessary

It should be noted that respondents do not necessarily traverse through each process when answering every question, nor is the list of provided processes exhaustive; rather, each process is a tool that respondents may consciously or unconsciously use when responding (Tourangeau et al., 2000). Exactly which set of processes they utilize will vary depending on the extent to which they want their answer to be accurate, how quickly they need to produce it, and other related factors. Descriptions of the individual

stages, as well as the response effects (i.e., measurement errors) that can emerge in each are provided below, all of which stem from Tourangeau and colleagues (2000).

Comprehension

Simply stated, the comprehension stage is that in which the respondent makes sense of the question and what it is asking him or her to do. It is in this stage that respondents interpret the focus of the question, the meaning of the response options, the instructions, and the definitions of key terms.

Response effects in this stage result from a misunderstanding of what a question is asking. Question wording and formatting play a key role in how a respondent interprets what is being asked, and misunderstandings may result from a variety of factors including a deficit in attention, failure to read instructions, double-barreled questions, inclusion of unfamiliar terms, and overcomplication. Characteristics of questions that minimize respondents' misunderstanding, and as such minimize error associated with this component, are provided later in this chapter.

Retrieval

The retrieval component of survey response involves recalling relevant information from long-term memory. In this component, respondents generate retrieval cues, recover memories, and fill in missing details. Aday and Cornelius (2006) describe several strategies that respondents may use to retrieve and recall events:

- Setting limits to their answers based on previous experiences or an implicit of implied comparison with others;
- Rounding estimates to prototypical values;
- Thinking of a kind of autobiographical time line or landmark public or private event as a point of reference;
- Disaggregating the task into a series of simpler tasks then summing or imputing a summary based on these discrete computations;

- Considering what they typically do or what they should do.

Again, there are a variety of response effects that emerge during this stage of the response process. Several characteristics of the recalled material and the questions can affect the accuracy and completeness of the retrieval, including the distinctiveness of the events to the respondent, the degree of fit between the terms used in the question and the event's original encoding, the number and quality of the cues that the question provides for retrieval, the source of the memory (direct or secondhand), and the length of time since the event occurred. For example, Wright, Gaskell, and O'Muircheartaigh (1998, as cited in Aday & Cornelius, 2006) found that adults were more likely to be able to recall an important event if it was associated with something they perceived to be important or to which they had emotional attachments.

Judgment

Because retrieval does not yield an explicit answer to many questions, respondents must often then assess the completeness of the memories they retrieve, draw inferences, and make estimates based on their partial recall. All of these processes are utilized in the judgment stage of Tourangeau's model. There are, at minimum, five major types of judgments that come into play when one is responding to a survey question:

- Judgments regarding the completeness or accuracy of the retrieval;
- Inferences based on the process of retrieval;
- Inferences that fill in gaps in what is recalled;
- Integration of the products of retrieval into a single overall judgment; and
- Estimates that adjust for omissions in retrieval.

The first three depend on the relation between judgment and retrieval and the others are used to transform retrieved information into an appropriate answer.

A variety of response errors may occur in the retrieval stage. The problems that specifically may lead to the overreporting or underreporting of health behaviors occur in situations when respondents (1) telescope, or include events from outside the time period being asked in the question, and (2) omit events that should have been included within the reference period (Aday & Cornelius, 2006). A respondent's ability to recall events is a function of the time period over which the events are to be remembered and of the salience or significance of the event to the person. The shorter the recall period, the less likely it is that respondents will omit events but the more likely they will telescope behaviors from the surrounding periods of time. Additional response effects may occur in this stage through the use of generalizations rather than specifics, and similarly, the combination of piecemeal information rather than complete pictures (Aday & Cornelius).

In addition to factual recall such as dates and the number of times someone engaged in a particular behavior, these effects are noticeable in answering questions about attitudes. A significant proportion of the population does not hold stable, crystallized views on a variety of constructs, and even when they do, these views may not lend themselves to a clear-cut response to a survey question. For such attitudinal questions, the judgments that the respondents make about the issue may be particularly critical as they think through how they will answer the question. Respondents may be more likely to have a strong response to an attitudinal question when the topic is important to them (Areni, Ferrell, & Wilcox, 1999, as cited in Aday & Cornelius, 2006). Ultimately, a variety of personal judgments and strategies are used in this stage to begin to formulate a question response, many of which contain unique sources of response error.

Response

The response component of this model is the stage during which respondents

finalize the response they will give to the survey question—if any. There are two types of

processes in the response stage: (1) mapping the answer onto the appropriate scale or

response option, and (2) editing the response for consistency, acceptability, and other

criteria. Even when respondents have a clear answer to report, there may be more than

one appropriate response (e.g., sometimes vs. often). Depending on a variety of factors,

respondents may deliberate to choose the best possible response or may be content to

select the first that applies.

Additional sources of error may also be introduced in this stage. Perhaps one of

the greatest response effects that occurs in this model component is that of social

desirability, or the tendency of respondents to distort answers in ways that will make

them look better or will avoid making them look bad (Fowler, 1995). In this phase of

Tourangeau's model, respondents weigh the benefits and risks of responding truthfully—

or of responding at all. Some individuals decide rationally if they want to answer

honestly, and others make this decision more subconsciously. In the case of the latter, the

accurate answer is not how the respondent wants to see him or herself (Fowler, 1995). To

reduce social desirability bias in survey research, which is a notable issue in health

surveys in particular (e.g., Di Iorio, 2005; Aday & Cornelius, 2006), researchers can (1)

assure confidentiality of responses and communicate effectively that protection is in

place; (2) communicate as clearly as possible the priority of response accuracy; and (3)

reduce the role of an interviewer (i.e., utilize paper-and-pencil, computer-assisted

technology, and e-surveying more frequently) in the data collection process (Fowler, 1995).

Summary

Tourangeau's Model of Survey Response provides a theoretical understanding for how people respond to questions on surveys. Yet, it is important to articulate that this is unlike stage models commonly cited in health behavior and health education. For many survey questions, survey takers follow the logical sequence specified in Table 2.3. However, this is not a necessary assumption of the model. Rather, evidence suggests that there are a large number of paths one may take to reach an answer to a survey question. Respondents can carry out components in parallel, backtrack through components previously visited, and/or completely skip components altogether. Tourangeau and colleagues (2000) state:

> "…[there are] quite a large number of paths to an answer are possible, depending on the effort that respondents are willing to invest and on the interplay between retrieval and judgment. In each case, the path traverses a subset of the process identified here – which may be carried out well or sloppily, in parallel or in sequence, and with or without backtracking – as circumstances and motivation dictate" (p. 19).

The last point in the quote above is of particular relevance. Though this model suggests a logical and comprehensive process through which people travel when answering survey questions, most respondents take an average of only 5 seconds to respond to a survey question (Tourangeau et al., 2000). Because surveys are typically low-stakes events in most respondents' lives (Fowler, 1995), survey designers may use the model above to create and administer questions that relieve respondent burden and reduce the respondent effects at each stage of this model, thereby increasing the reliability and validity of the collected data. Characteristics

of what generally constitutes these well-designed survey questions are presented in the next section of this chapter.

Characteristics of Reliable and Valid Survey Questions

There is an almost infinite body of desirable and useful information that can be gathered from people simply by asking them questions. What, then, makes a question worth asking? It has been suggested that "a good question is one that produces answers that are reliable and valid measures of something we want to describe" (Fowler, 1995, p. 2). The combination of expanded theory development and increased survey administration in the late 20th century contributed to a great deal of research into characteristics of questions that ease respondent burden and produce these aforementioned reliable and valid survey data (Marsh, 1982). Because they are the vehicle of data collection, the questions themselves are perhaps the most critical components in achieving high quality data from a survey. The best of sampling schemes and estimation strategies will not yield accurate data if the answers provided by the respondent are not meaningful (DeMaio, Rothgeb, & Hess, 1998).

Even small modifications to survey questions can produce dramatic differences in results (Higher Education Research Institute [HERI], n.d.), and a variety of factors influences how people answer questions such as question wording, the wording of response alternatives, the order of response alternatives, the mode of data collection, and contextual effects (Fowler, 1995). In his book, *Improving Survey Questions*, Fowler (1995) states there are five basic characteristics of questions and answers that are fundamental to a good measurement process:

- Questions need to be consistently understood;
- Questions need to be consistently administered or communicated to respondents;
- What constitutes an adequate answer should be consistently communicated;
- Unless measuring knowledge is the goal of the question, all respondents should have access to the information needed to answer the question accurately; and
- Respondents must be willing to provide the answers called for in the question.

Even when psychometric indicators are acceptable, respondents may interpret some items in disparate or unintended ways (Ouimet et al., 2004). As such, there are a variety of challenges of writing "good" survey questions, a list of which is provided in Figure 2.2.

Challenges of Reliable and Valid Survey Question Writing

1. Defining objectives and specifying the kind of answers needed to meet the objectives of the question.

2. Ensuring that all respondents have a shared, common understanding of the meaning of the question. Specifically, all respondents should have the same understanding of the key terms of the question, and their understanding of those terms should be the same as that intended by the person writing the question.

3. Ensuring that people are asked questions to which they know the answers. Barriers to knowing the answers can take at least three forms:
   a) Never having the information needed to answer the question
   b) Having the information at some point, but being unable to recall the information accurately or in the detail required by the question
   c) Difficulty in accurately placing events in time

4. Asking questions that respondents are able to answer in the terms required by the question. It is possible to ask questions to which respondents know the answers but are unable to answer the way that investigators want because of a lack of fit between the desires of the investigator and the reality about which the respondent is reporting.

5. Asking questions respondents are willing to answer accurately.

Figure 2.2. Challenges of writing survey questions as described by Fowler (1995, p. 9).

To address some of these challenges, survey methodologists have developed a list of item writing guidelines they believe reduce bias. These guidelines, many of which are intuitive, are summarized from Di Iorio (2005) and Fowler (1995), and include:

- Write items that are related to the purpose of the survey and that address the objectives;
- Write each item as a complete sentence with a subject and a verb;
- Write items that are short and concise when requesting information about neutral topics;
- Write open-ended or long items when requesting information about sensitive topics;
- Write response choices that are mutually exclusive and exhaustive;
- Spell out acronyms;
- Define unusual terms;
- Write items that contain only one idea;
- Write items that are specific;
- Use simple words;
- Highlight terms that could be easily missed or that are particularly important;
- For recall items, allow a specific time period that is consistent with the behavior;
- Minimize the difficulty of recall and reporting tasks given to respondents;
- Give respondents help with recall and placing events in time by encouraging the use of association and other memory aids;
- Use ranges rather than precise values for sensitive items;
- Use response options such as "Don't know" and "Not applicable" sparingly;
- Make careful decisions about which response options to include;
- For objectives that pose special definitional or are recall challenges, use multiple questions;
- Make sure the form of the answer to be given fits the reality to be described;
- Design all aspects of the data collection to minimize the possibility that any respondent will feel his or her interests will be best served by giving an inaccurate answer to a question.

Correspondingly, researchers in this field have also suggested it is best to avoid the following when writing surveys: questions with the word "not", questions that contain "jargon" or regional expressions, ambiguous questions, and questions containing value-laden or biased words (Di Iorio, 2005).

As the field of survey research continues to grow, particularly with the increase in web-based surveying, the empirical evaluation of survey questions remains critically important (Fowler, 1995). Failure to investigate the interpretation of questionnaire items may result in misinterpretation (by respondents and researchers), falsified answers, missing responses (to the questionnaire as a whole or to a particular item), and potentially offended respondents (Bowden et al., 2002). Hence, survey researchers have designed a

variety of techniques to best assess how well the questions they ask meet the process standards and questionnaire objectives, including (1) focus group discussions; (2) cognitive interviews, in which people's comprehension of questions and how they go about answering questions is probed and evaluated; and (3) field pre-tests under realistic conditions (Fowler, 1995). As the focus of this dissertation is a field pre-test of a revised survey instrument, this technique is explored in greater depth later in this chapter. First, though, because the ultimate psychometric "standards of goodness" of survey items, questions, and scales are the validity and reliability with which they produce answers that measure something (Aday & Cornelius, 2006), current literature reviewing these concepts and the methods though which they are established in survey research is provided below.

Issues and Practices in Establishing Reliability and Validity in Survey Research

Survey research and psychometrics are part of the larger field of measurement, the conceptualization of which is depicted in Figure 2.3. Measurement has been defined as "the assignment of numerals to aspects of objects or events according to rule[s]" (Stevens, 1959, p. 24). Perhaps the two most fundamental concepts of measurement are reliability and validity. *Reliability* refers to the accuracy of a given measurement, and a reliable survey question is one that yields similar results when administered repeatedly to similar samples or populations (HERI, n.d.). *Validity*, by contrast, refers to whether a survey question actually taps into the true underlying concept it attempts to measure—in other words, the legitimacy of the attained scores as a measure of the chosen attribute (HERI; Di Iorio, 2005). No measure ever produces perfectly reliable or valid scores, and the extent to which a measure deviates from the ideal level of reliability and validity is

referred to as *measurement error*. "Error is an inevitable but undesirable component of measurement. Error affects the reliability and validity of scores, and, if severe, can lead to flawed interpretation of the data" (Di Iorio, p. 105).
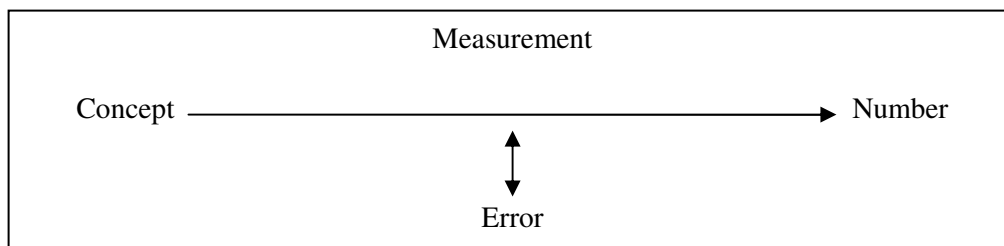
Figure 2.3. Conceptualization of the measurement process.

Two types of error exist: random and systematic. Random, or unsystematic, error impacts reliability and does not affect the attribute being measured in the same way each time a measurement is taken. On the other hand, systematic error affects measurement in the same direction each time and hence affects the validity of measurement. The following summary presents a brief overview of factors that can produce these errors, methods to correct these errors (see Table 2.4 for a very basic synopsis), and methods by which reliability and validity can be assessed.

Assessing reliability in survey research

As mentioned above, reliability refers to the consistency and stability of measurement and is impacted by random measurement error. A certain amount of random error will always be present in measurement; as such, the assessment of reliability of survey indicators is not to determine whether they possess random error, but rather, the extent to which they possess this error (Carmines & Zeller, 1979). According

Table 2.4. Sources and solutions of survey measurement. Adapted from Aday & Cornelius (2006).

Systematic errors: Low or poor validity

Establish:	*By monitoring and evaluating….*
• Content validity	the systematic departures in the content of a survey question from the meaning of the concept itself
• Criterion validity	the accuracy of the answers based on comparisons from another data source
• Construct validity	the strength of hypothesized relationships of the concept being measured with other measures or concepts

Variable errors: Low or poor reliability

Establish:	*By monitoring and evaluating….*
• Test-retest reliability	random variation in answers to a survey question due to when it is asked
• Inter-rater reliability	random variation in answers to a survey question based on who asked it
• Internal consistency reliability	random variation as one of a number of questions asked to construct a summary scale

to classical test theory, these random disturbances sometimes occur in the positive direction and at others occur in the negative direction; the magnitudes of these positive and negative errors are similar and subsequently balance one another out (Carmines & Zeller). Random errors that affect survey measurement may stem from a variety of factors, many of which are transient, such as respondents' mood, temperature of the room, and time of day at which the survey is administered (Trochim, 2006).

Several different methods to assess reliability are available, and the selection of reliability procedures depends on a number of factors, including the attribute being measured, the type of instrument, the investigator's skills and available time, the availability of research participants, and data collection time and efforts (Di Iorio, 2005). The four most commonly used forms of reliability assessment include inter-rater reliability, test-retest reliability, parallel-forms reliability, and internal consistency

reliability (Trochim, 2006). Briefly, inter-rater reliability is used to assess the degree to which different raters give consistent estimates of the same phenomenon; test-retest reliability assesses the consistency of a measure from one time to another; parallel forms reliability assesses the consistency of the results of two instruments constructed in the same way from the same content domain; and internal consistency reliability assesses the consistency of results across items within a survey, usually from a scale (Trochim).

Internal consistency reliability estimates are perhaps the most commonly used in health survey research (Di Iorio, 2005). These approaches, which require only a single administration of a survey, estimate reliability from the assessment of correlations and relationships among items within a scale; accordingly, a necessary criterion for establishing internal consistency reliability is the measure must consist of more than one item (Di Iorio). The reliability coefficient is often assessed by calculating the coefficient alpha (Cronbach, 1951), a process through which each item is correlated with each other item on a scale and alpha is computed by taking the average of individual item-to-item correlations and adjusting for the number of items (Di Iorio). Higher coefficient alphas are associated with scales that have more items and higher interitem correlations, and it has been suggested that an alpha coefficient of .70 or greater demonstrates sufficient reliability (Nunnally & Bernstein, 1994, as cited in Di Iorio, 2005) of a survey scale.

Assessing validity in survey research

The validity of survey questions refers to the degree to which responses differ systematically from (1) the meaning of the concept they were intended to measure, (2) related questions about the same concept, and (3) theories or hypotheses about their relationships to other concepts (Aday and Cornelius, 2006). Broadly speaking, the term

validity refers to the legitimacy of an instrument—that is, does it measure what it intends to measure? More specifically, validity is the degree of support obtained for the *interpretations of scores* on an instrument when the instrument is used for its *intended purpose* (Di Iorio, 2005).

Assessing evidence of validity in survey research is often quite cumbersome, particularly when evaluating the validity of self-reports of behaviors, attitudes, and psychosocial states (Di Iorio, 2005). There are three common types of validity for which researchers attempt to demonstrate evidence in health survey research: content, criterion, and construct. A concise overview of each and the methods used to assess it are provided below, the summary of which generally comes from Aday and Cornelius (2006).

Content validity refers to the extent the chosen questions represent the concepts they are intended to reflect. Therefore, content validity is an issue of item sampling adequacy—that is, the degree to which a specific set of items reflects its content domain. It is thus easiest to evaluate evidence of content-related validity when the domain to be assessed is well defined (DeVellis, 2003). Methods commonly used to assess evidence of content validity are:

- Reviewing the literature to determine the scope of the concepts to be measured;
- Using questions and variables on the same topic that have been used in other studies; and
- Asking an expert panel whether, in their judgment, the questions represent the concept being measured.

Criterion-related validity refers to the extent to which the data collected from a survey agrees with some "true" value or "gold standard" of measurement of a construct. Whether the theoretical bases for these associations are understood is irrelevant to criterion-related validity; therefore, criterion-related validity is often more of a practical issue than a scientific one (DeVellis, 2003). It has been suggested that criterion-related

validity can be divided into two subtypes: concurrent (demonstrating a measure relates to other criteria simultaneously) and predictive (demonstrating a measure can predict future concrete events).

A common technique used to establish evidence of concurrent validity is to examine the relationship between the items of interest and other surveys and/or measures that have been previously validated. For example, when researchers correlate respondents' self-report of a variety of diagnoses over the last 12 months with medical records from the same time period, they are assessing evidence of concurrent validity.

Evidence of predictive validity stems from the ability of survey items' ability to predict expected future events. This type of evidence is typically demonstrated in longitudinal studies or in experimental and quasi-experimental study designs employed to reflect changes in respondents' behaviors, attitudes, or knowledge over time.

Criterion validity was, for many years, the mainstay of validity assessment (Di Iorio, 2005). However, in many cases, researchers realized there were often no suitable criteria for comparison, a phenomena that was particularly salient for the measurement of latent psychological variables. To address this problem, researchers developed strategies to evaluate construct validity. Because constructs rarely exist in isolation, survey methodologists suggested that if an instrument measured what it claimed to measure, data from the instrument would exhibit patterns that would be expected according to theory (Di Iorio). Hence, as articulated by (Aday & Cornelius, 2006):

> "Evaluations of the construct validity of a survey variable assume that there are well-developed theories or hypotheses about the relationships of that variable to others being measured in the study. Construct validity examines whether and how many of the relationships predicted by these theories or hypotheses are empirically borne out when the data are analyzed. The more often these hypothesized relationships are confirmed, the greater the construct validity of the survey variables is assumed to be" (p. 66).

Typically, evidence of construct-related validity is demonstrated through the use of statistical models and analyses. Discriminant and convergent validities are frequently examined together and involve studying the strengths or patterns among data collected from a variety of variables (Spector, 1992). Convergent validity refers to a strong correlation between different measures of the same construct. Discriminant validity suggests measures of different constructs should relate only modestly with one another (Spector). Hence, a multitude of strategies to evaluate construct-related evidence of validity may be used. These include:

- Testing hypotheses in regards to the level to which various constructs correlate positively or negatively with each other;
- Evaluating the levels to which groups known to differ on a particular construct (e.g., men and women with stress levels) actually differ in the data collected;
- Examining the ways scores on an instrument change over time in instances where changes are expected to occur; and
- Developing a multitrait-multimethod matrix to assess levels of convergent and discriminant validity.

As evidenced in this brief overview, there are numerous strategies that may be used to establish evidence of content, criterion, and construct validity of a survey. Ultimately, there are several factors that contribute to a person's score on a scale or other assessment, including characteristics of the instrument, characteristics of the person, and characteristics of the setting in which the survey is administered (see Table 2.5 for a list of common respondent and instrument errors that may impact measurement validity). Therefore, it is not surprising that the literature suggests,

> "…validity must be assessed among different groups and in different settings… Validity is not an all-or-none principle; rather, it is an evolving property. New findings may either enhance validity or detract from validity for a particular group of respondents. Evidence is never complete; thus the process of validation is continual" (Di Iorio, 2005, p. 236).

Though they are often cumbersome, the literature suggests a crucial need for validation

studies of survey instruments, particularly in the health field (e.g., Aday & Cornelius,

2006; Di Iorio, 2005). Subsequently, this dissertation, which will examine evidence of

criterion- and construct-related validity of a health survey will fufill this critical research

need for validation of large-scale surveys.

Table 2.5. Respondent errors and instrument factors commonly affecting measurement
validity in survey research. Adapted from Di Iorio (2005).

Respondent errors	Instrument factors
- Over- or underreporting agreement with items - Over- or underreporting frequency of events - Over- or underreporting positive or negative attitudes - Providing partial answers - Providing incorrect answers - Making recording errors	- Inadequate or incorrect instructions - Poor formatting - Illogical order of items - Use of vague or unfamiliar terms - Response options that vary between items - Items that fail to correspond to the construct measured by a scale - Response options that fail to fit the question stem

Summary

Both systematic and random errors are inherent to the measurement process, and

as such are expected. No survey or scale is perfectly reliable and valid, but it is important

that survey developers utilize strategies to increase the extent to which their

measurements minimize error and maximize "truth" in the constructs they assess. The

literature provides a variety of methods through which evidence of reliability and validity

can be demonstrated. As stated above, the process of minimizing errors and

understanding survey responses is iterative and extensive, and not surprisingly, is a part

of every stage of the survey development process. Field pre-testing is one strategy that

may be used to establish preliminary evidence of reliability and validity before surveys

are implemented in a large-scale basis. This method is described in detail in the next section of this chapter.

Issues and Practices in Pre-testing Survey Questionnaires

Though they date back to as early as the 1930s (Presser et al., 2004), survey practitioners have significantly increased their use of an evolving set of questionnaire pre-testing methods over the last 15 years in an attempt to collect more reliable and valid data (Rothgeb, Willis, & Forsyth, 2001). The following summarizes basic principles and challenges of these methods, continues with their specific use in survey revisions processes, and provides an example of a recent study that used experimental pre-testing strategies as part of a survey revisions process.

Overview of survey pre-testing

Pre-testing, the utility of which is documented extensively in the literature (e.g., Rothgeb et al., 2001; Presser et al., 2004; Presser & Blair, 1994; Ouimet et al., 2004; Bowden et al., 2002; De Maio et al., 1998; Marsh, 1982; Fowler, 2004; Moore, Pascale, Doyle, Chan, & Griffiths, 2004; Tourangeau, 2004) encompasses a range of activities, including evaluating (1) individual questions, (2) the questionnaire as a whole, (3) the feasibility of sampling and data collection procedures, and (4) the procedures for coding and computerizing the data (Aday & Cornelius, 2006). A pre-test of a survey is essentially its "dress rehearsal" among a population and in a setting similar to that which will be used in its ultimate implementation (Presser et al., 2004). In general, the rationale for pre-testing is that trying out a questionnaire with respondents will reveal problems that even the most experienced survey researcher cannot diagnose (Presser & Blair,

1994), and that it provides an opportunity for researchers to gauge the meaning attributed to survey questions before substantial investments are made in the wrong questions or in questions where the researcher cannot be sure about what precisely is being asked (Bowden et al., 2002).

Interestingly, though research suggests no survey should ever go into the field without a trial run of the questionnaire and data collection procedures to be used in the final study (Aday & Cornelius, 2006), many large scale survey efforts fail to pre-test the instrument prior to implementation or do so inadequately (Rothgeb et al., 2001; Presser et al., 2004; Tourangeau, 2004). Numerous reasons have been suggested for this breach of suggested practice, including lack of consensus and guidance as to appropriate methodology (Bowden et al., 2002; Marsh, 1982), acknowledgement within the field that conventional methods may fail to uncover certain problems (Rothgeb et al., 2001; Presser et al., 2004), lack of published studies with pre-test results (Presser et al., 2004), and lack of resources (e.g., time, subjects, money) needed to conduct a meaningful pre-test (Tourangeau, 2004). It has additionally been suggested that even when pre-tests have the ability to uncover problems with survey items, they then fail to provide the means by which the problems may be fixed (Presser & Blair, 1994; Fowler, 2004).

Historically, the most common methods through which pre-tests identified problems with questions were (1) interviewer feedback from their preliminary experiences administering questionnaires in the field (DeMaio et al., 1998), and (2) response analysis, in which questions were assessed for the amount of "don't know" responses or refusals that emerged in response tallies (Presser et al., 2004; Singleton & Straits, 1999). Though these methods may have provided some important information,

they were utilized primarily for in-person or telephone-based survey interviews (Presser

et al., 2004), which were more common in the mid- to late-20th century before the

emergence of computer-assisted and web-based survey technologies.

As stated above, current literature suggests a renaissance of pre-testing as an

imperative step in the survey development process over the last fifteen years, largely due

to a paradigm shift in the goals of these techniques as well as changing needs to assess

computer- and internet-based surveys. Once utilized only to identify drastic challenges

experienced by interviewers and respondents, pre-testing today is more commonly used

to assess the more broad concerns of improving data quality (i.e., reliability and validity)

so measurements best meet survey objectives (Presser et al., 2004). Simultaneously,

newer pre-testing methods have emerged as theory and practice have advanced in the

field, all of which use information from respondents to modify and evaluate survey

questions and design. These include cognitive interviews, behavior coding, response

latency, vignette analysis, formal respondent debriefings, experiments, and statistical

modeling (Presser et al., 2004). Definitions of these methods and basic strategies they

employ are summarized in Table 2.6.

Each of the methods above provides a unique capacity to identify questionnaire

problems, and the most robust information comes from the triangulation of information

derived from a combination of their use (Ouimet et al., 2004). In addition to providing an

understanding as to how respondents answer survey questions, these strategies also

provide information as to the problems that may exist with particular questions. Once

problems or challenges with survey questions are identified, items are frequently revised

in an effort to address the challenges (Presser et al., 2004). Further, surveys commonly

Table 2.6. Definitions and strategies of various pre-test methods in the field of survey research.

Pre-test method	Definition	Basic strategies used with method
Cognitive interview (Conrad & Blair, 2004)	A technique that assumes respondents accurately verbalize their experiences and asks them to think aloud to verbally report the process through which they go when answering survey questions.	Researchers ask questions (either concurrently or retrospectively) of respondents one-on-one to discern their experiences when answering questions on surveys; respondent feedback helps to identify problematic questions.
Behavior coding (Presser et al., 2004)	A technique that monitors interviews or reviews tapes of interviews for a subset of the interviewer's or respondent's verbal behavior in the question and answer interactions to discern problematic questions.	Problem questions are identified when high frequencies of certain behaviors (e.g., interviewer failing to read question verbatim, respondent asking for clarification) are present.
Response latency (Presser et al., 2004)	A technique that identifies the amount of time it takes for respondents to answer a survey question (i.e., the amount of time between when a respondent finishes reading a question and when he or she enters a response).	Computer-assisted technology is used to evaluate the amount of time to produce a response; longer delays are commonly believed to indicate respondent uncertainty and possible question problems.
Vignette analysis (Martin, 2004)	A technique in which respondents are presented with hypothetical scenarios to evaluate the ways by which they reason and respond to a variety of situations.	Presenting respondents with hypothetical scenarios allows researchers to explore conceptual domains, test consistency of respondents' interpretations with survey intents, evaluate alternative questionnaires, and analyze concept dimensionality.
Formal respondent debriefing (Martin, 2004)	A technique that asks respondents general, probing questions or standardized, retrospective questions after survey items to assess their experiences and reasoning when answering the survey items.	Question-testing methods and "double interviews" (i.e., asking questions after survey items to gather why respondents answered a certain way) identify problems in respondents' understanding and comprehension of survey items, interpretations, subjective reactions or thoughts, and direct measures of missed/misreported information.
Experiments (Fowler, 2004)	A technique in which respondents are randomized to receive different versions of questions or survey instruments to determine differences in survey responses and/or respondent experiences.	Results of different questions or instruments may be analyzed to determine differences in response distributions, validation against a standard, and usability.
Statistical modeling (Presser et al., 2004)	A technique in which data collected from surveys are analyzed using a variety of statistical techniques to evaluate relationships between variables and evaluate measurement properties.	Quantitative analyses such as latent class models, multitrait-multimethod techniques, confirmatory factor analysis, and structural equation modeling are utilized to assess measurement and construct characteristics as well as evidence of reliability and validity.

undergo revisions to improve their measurement capabilities in response to respondent,
stakeholder, and expert feedback, and it is important to assess the extent to which these
revisions constitute true improvements in survey design (Presser et al., 2004). Pre-testing
is frequently used as an evaluative tool in survey revisions, and its use is described below
with a special focus on experimental designs.

Pre-testing as part of the survey revisions process

As stated earlier in this chapter, even small changes to survey questions can
produce very significant differences in item responses (HERI, n.d.). Surveys may
undergo revisions for a variety of purposes, many of which have been previously
discussed. Frequently, surveys are revised when questions are thought to be out-of-date
or psychometrically flawed in some way (Fowler, 1995). Once these changes have been
made, it is important to evaluate the impact of the edits on survey results. Survey
developers commonly utilize many of the qualitative strategies listed above (e.g., focus
groups, cognitive interviews) to assess how respondents interpret and answer the edited
questions (Presser & Blair, 1994; Ouimet et al., 2004; Moore et al., 2004; Tourangeau,
2004), but when used alone the qualitative methods are insufficient in the information
they provide. The importance of quantitative evaluation methods to evaluate survey
revisions cannot be underscored (Presser et al., 2004; Presser & Blair, 1994; Moore et al.,
2004; Tourangeau, 2004; Fowler, 2004); Fowler (2004, p. 174) cites three pivotal reasons
for quantitative revisions analysis:

1. Sometimes "fixing" a problem makes a question worse from other perspectives.
   It is important to know whether the effects of the fix constitute an improvement
   over the original from the various perspectives from which questions must be
   evaluated.

2. One of the most important conservative forces for keeping problem questions is the desire to use questions that have been used in previous surveys. Reviewers often like the fact that an item has been used previously, whether or not it has been carefully evaluated. When problems are found is questions with a pedigree, how much the problems affect data quality and how improved versions of the question will affect estimates are important considerations.

3. When a researcher "fixes" a problem, it is highly valuable to know whether or not the new question produces data that are more likely to be valid or better by some standard that the original.

Though it is utilized less in practice than in theory due to a lack of knowledge, time, subjects, and other resources, the experimental pre-test is perhaps the most commonly suggested method for evaluating differences between original and revised survey questions (e.g., Presser et al., 2004; Presser & Blair, 1994; Moore et al., 2004; Tourangeau, 2004; Fowler, 2004). The premise of an experimental pre-test is simple: two comparable samples of respondents are asked different versions of questions designed to achieve the same question objective (Fowler, 2004). A fundamental characteristic of an experimental field pre-test is that respondents are randomly assigned to alternative questionnaire versions under controlled conditions (Moore et al., 2004).

Once data are collected from respondents on both versions of the survey, the questions may be compared in a variety of ways (Fowler, 2004): (1) Researchers may compare the response distributions; if they appear virtually identical, one may conclude that the two questions are virtually the same question; (2) Researchers may attempt to compare the validity of the two questions using a standard against which to assess answers; results may tell which answers are "more valid"; (3) Researchers may compare the "usability" of the two questions by examining the percentages of missing responses or "don't knows"; these results may inform researchers as to which version of the question introduces less respondent burden and confusion.

Though there are few controlled studies documenting experimental analyses of survey revisions pre-tests, the potential of this method to provide unique psychometric information is unwavering in the literature. The following is a brief summary of a recent experiment used to assess the impact of question and instrument revision on responses.

Example of an experimental pre-test study to evaluate survey revisions

Recent literature provides a summary of results from an evaluation of instrument modifications on a telephone survey of adults enrolled in health plans in Washington State (Fowler, 2004). In this study, respondents were randomly assigned to receive either the original (final n = 261) or alternative version (final n =299) of a question asking about their ability to get appointments for routine care. The study reported the proportion of respondents reporting they were "always" able to get appointments differed significantly between the two versions of the question. Table 2.7 summarizes the study findings.

Table 2.7. Results of an experimental pre-test for one survey question. Summarized from Fowler (2004).

Original Question	Revised Question	Results
In the last 12 months, how often did you get an appointment for *regular or routine* health care as soon as you wanted—always, usually, sometimes, never? ("Always" coded to "Yes")	In the last 12 months, were you always able to get an appointment as soon as you wanted?	Differences in proportions were statistically significant (p<0.001) across versions of the survey
Always Got Appointment (n = 261) • Yes 47% • No 53%	Always Got Appointment (n = 299) • Yes 66% • No 34%	

These results indicate that even minor changes in question wording can produce highly significant differences in responses. However, the study investigator makes no claims as to which item is a more valid indicator of the ability of the subjects to make routine care

appointments. As such, no real conclusions other than "these items are different" can be made (Fowler, 2004).

The literature provides further examples of experimental pre-tests, but, like the example above, most are conducted with interviewer-administered surveys (e.g., Moore et al., 2004; Barker, Gfroerer, Casper, & Lessler, 1998) as opposed to e-surveys and as such have limited applicability to this study. Or, when they evaluate e-surveys, the experiment is used to evaluate the impact of using web-based surveys as compared to pencil-and-paper based surveys (e.g., Whitehead, 2007; Parks, Pardi, & Bradizza, 2006) Thus, this study will add immensely to the body of literature on experimental field pre-tests examining the effects of revisions of questions asked on web-based surveys.

Summary

Survey pre-testing allows survey developers to preliminarily assess the experiences of respondents with their instrument. Pre-testing, and particularly field experiments, have tremendous utility when examining the impact of question and survey revisions. Presser and colleagues (2004) suggest, "To determine whether the revisions are improvements… there is no substitute for experimental comparisons of the original and revised items" (p. 118). Yet, despite their usefulness, pre-tests in general and particularly experimental designs are not commonly part of the presurvey question evaluation process (Fowler, 2004). This may be especially true among instruments assessing college student experiences (Ouimet et al., 2004). Data from pre-test studies have the potential to contribute significantly to the theory and practice of survey research by identifying the extent to which survey responses are affected by various problems (Presser & Blair, 1994), by providing preliminary information about the validity of survey revisions

(Tourangeau, 2004; Presser et al., 2004), and when applicable, by helping to decide which of two alternative forms of a question is most psychometrically sound (Fowler, 2004). This dissertation, which will use experimental field pre-testing methods to assess the impact of revisions to the mental health indicators of the ACHA-National College Health Assessment will fill a critical research gap in this field.

Conclusion

Evidence suggests the prevalence of mental health disorders and challenges is increasing in today's college students, and they are at particular risk for anxiety, depression, eating disorder, deliberate self-harm, substance abuse, and suicide. Correspondingly, the collection of reliable and valid data for these phenomena remains a critical college and public health concern. Large-scale online surveys are common and appropriate data sources for health constructs among college students; however, the literature and experts in the field have identified numerous gaps in the ACHA-NHCA survey, the only current large-scale online survey measuring college student mental health.

Surveys are often revised when such gaps in content and/or psychometric challenges are identified, and the importance of evaluating the impact of these revisions and the extent to which revised items are reliable and valid cannot be understated. Experimental field pre-tests, in which respondents similar to the target audience are randomized to participate in either the original or revised survey data collection efforts to discern differences in response patterns, are considered the quantitative gold standard in

these types of evaluations; yet, they are often underutilized due to lack of knowledge or resources.

This study, which evaluates the results of an experimental field pre-test of the revised mental health indicators on the ACHA-NCHA contributes greatly to a variety of bodies of literature, including (1) preliminary estimates of the prevalence of a variety of mental health concerns in the college student population; (2) evidence of reliability and validity of these mental health indicators in a college population; (3) the impact of how modifying question format and survey design impacts responses between two versions of a comparable web-based survey measuring student mental health; and (4) the issues and challenges in quantitative analysis of data collected from experimental field pre-tests of a revised large-scale survey.

CHAPTER III: METHODS

Introduction

Surveys continue to be an essential tool through which researchers gather information about a variety of health phenomena. The American College Health Association-National College Health Assessment (ACHA-NCHA) is a large-scale health survey monitoring a variety of health constructs of students attending PSIs in North America. As is common to a number of large-scale survey efforts, the ACHA-NCHA underwent an extensive revisions process beginning in 2005 as a result of numerous feedback mechanisms. During this revisions process, the section of the survey measuring college student mental health was expanded and modified dramatically, largely because of mounting evidence suggesting the prevalence of a variety of mental health concerns has been increasing on college campuses over the past few decades (e.g., Sharkin, 2006; Gallagher, 2007; Gilbert, 1992; Kadison & DiGeronimo, 2004; Kadison, 2006), and because there is a documented need for such data (e.g., Soet & Sevig, 2006).

The ACHA utilized numerous strategies described as appropriate and relevant in the literature when revising the survey, including expert review, reviews of the literature, solicited feedback from respondents, and focus groups (Presser et al., 2004; Conrad & Blair, 2004; Fowler, 2004; Fowler, 1995; Trochim, 2006; DeVellis, 2003; Singlteon & Straits, 1999). An overview of all revisions steps is included in Chapter One. As a final step before the survey was fully implemented with its target population, the ACHA pre-tested the revised survey under realistic field conditions in Spring 2007. Procedures included an experimental field pre-test which, as described in Chapter Two, is often cited

as the "gold standard" in the area of survey research as it allows researchers to determine how changes to survey questions produce changes in survey responses and subsequently impact the measurement quality of the data collected. Seven PSIs self-selected to participate in the pre-testing efforts, and students from each institution were randomized to receive either the original or revised surveys (see Appendices B.1 and B.2 for full versions of each instrument). This research and analyses, which focus exclusively on the mental health indicators of the ACHA-NCHA, has two broad purposes: (1) to determine how changing question wording and/or response options produced changes in students responses across versions of the instrument, and (2) to begin to demonstrate evidence of the reliability and validity of data collected from the mental health indicators on the revised ACHA-NCHA.

Research Questions

This study evaluated the following research questions and hypotheses using data collected during field pre-testing of the revised ACHA-NCHA:

- **Research Question 1:** What patterns of student responses emerge on each of the revised ACHA-NCHA mental health indicators?

    Research Question 1A: How do student endorsements of each item vary according to demographic characteristics such as gender, year in school, and race/ethnicity?

- **Research Question 2:** What differences exist in student response patterns between comparable original ACHA-NCHA mental health items and revised ACHA-NCHA mental health items?

Hypothesis 2: Changes in question formatting and response options between the original and revised ACHA-NCHA mental health indicators will result in significant differences in patterns of student responses on comparable items across versions of the survey.

- **Research Question 3:** To what extent are the mental health indicators on the revised ACHA-NCHA reliable?

  Hypothesis 3A: Measurement scales on the revised ACHA-NCHA will demonstrate sufficient levels of internal consistency in the pre-test sample.

  Hypothesis 3B: Reliability will not increase if any indicators are removed from measurement scales on the revised ACHA-NCHA in the pre-test sample.

- **Research Question 4:** To what extent are the mental health indicators on the revised ACHA-NCHA valid?

  Hypothesis 4A: As a result of comparison with data collected from other surveys of mental health among college-aged students, the mental health indicators on the revised ACHA-NCHA will demonstrate evidence of criterion-related validity in the pre-test sample.

  Hypothesis 4B: As a result of analyses demonstrating expected theoretical relationships between variables, mental health indicators on the revised ACHA-NCHA will demonstrate evidence of construct-related validity in the pre-test sample.

- **Research Question 5:** How do mental health indicators on the revised
  ACHA-NCHA differ from those on the original ACHA-NCHA in terms of
  their reliability and validity?

  *Hypothesis 5***:** Mental health indicators on the revised ACHA-NCHA will
  demonstrate evidence of higher levels of reliability and validity than those
  on the original ACHA-NCHA.

Study Design

This study is a reliability and validity analysis of a subset of cross-sectional data collected from the American College Health Association in an experimental field pre-test of the revised ACHA-NCHA from February through May 2007. Accordingly, this study's aims are to evaluate the psychometric properties of the revised mental health indicators using statistical analyses and comparisons to other national surveys when applicable. The background for how data were initially collected by the ACHA is provided below to highlight how experimental pre-test methods can be used to collect and evaluate data.

Seven PSIs self-selected to participate in ACHA's experimental field pre-test data collection efforts, and students from each institution—either all students from the institution or a random sample of students—were randomized to be invited to participate in either the original or the revised instrument. Students ultimately self-selected to respond to the surveys and self-reported on each question asked on the survey.

The initial data collection effort is considered an *experimental pre-test* because two comparable samples of respondents were randomized to respond to different versions

of questions designed to achieve the same question objective under controlled conditions (Fowler, 2004; Moore et al., 2004). For each institution that participated, these controlled conditions included:

- Email letters of invitation used to recruit participants were identical across versions at each PSI;
- Subject lines of emails used to recruit participants were identical across versions at each PSI;
- The length of time during which the survey was open for students to participate was identical across versions at each PSI;
- The incentives used to encourage participation were identical across versions at each PSI; and
- The number of contacts of non-responders and timing of non-responder contact was identical across versions at each PSI.

As mentioned in the introduction to this chapter, the purpose of this study was two-fold. First, data collected from the revised survey were compared to data collected from the original survey to determine how changes in question format and wording impact the ways in which students respond to the items. Second, data collected from the revised survey were assessed for evidence of internal consistency reliability, construct-related validity and criterion-related validity.

Study Population

The population of interest in this study is students enrolled in 4-year postsecondary institutions (i.e., non-community colleges) in the United States (i.e., U.S. college and university students). According to the National Center for Education Statistics Spring 2006 Integrated Postsecondary Education Data System (IPEDS) database, more than 17.9 million students enrolled in 6,607 PSIs in Fall 2005 (Knapp et al., 2007). Nearly two-thirds of all U.S. college and university students (61 percent) are undergraduates enrolled in 4-year institutions, which equates to approximately 10.9

millions students. In addition, approximately 3 million graduate students are enrolled in 4-year institutions each year.

Nearly half of 4-year PSI undergraduate enrollees (46 percent) are enrolled in their first or second year and 54 percent are enrolled in their third year or higher (U.S. Census Bureau, 2006). Further, students are very racially and ethnically diverse: in 2005, nearly a quarter of students in 4-year PSIs were a racial or ethnic minority. The undergraduate population at 4-year PSIs is 43.9 percent male and 56.1 percent female. The graduate population (including graduate students who attend professional schools in addition to those attending strictly 4-year institutions) is 40.8 percent male and 59.2 percent female (U.S. Census Bureau). An overview of the participating institutions that comprise the study population and their characteristics is provided below.

Study Participants

Participants for this study were students enrolled in seven (7) PSIs that self-selected to participate in the experimental field pre-test of the ACHA-NCHA. Characteristics of the participating institutions, the procedures used to recruit their students for participation in the survey efforts, and the characteristics of the two final student samples (original and revised instrument) are provided in the following sections of this chapter.

Overview of participating institutions

Between January and May 2007, 113 PSIs self-selected to participate in the ACHA-NCHA Spring 2007 survey efforts. As part of the self-selection process, PSIs pay a fee for each survey distributed to their students, a fee for each record of data processed,

and a flat fee for a report package. ACHA institutional members receive discount pricing (see Appendix A for more information as to types of membership available), and nearly all participating PSIs are ACHA institutional members. Of the 113 institutions that participated in Spring 2007 data collection efforts, 107 utilized random sampling or census (i.e., generalizable) techniques and were hence included in the ACHA-NCHA reference group, an aggregate sample with data from 71,860 students.

In order to conduct an experimental field pre-test with significant power to detect differences between original and revised survey questions, the ACHA-NCHA revisions committee and ACHA researchers decided to recruit a minimum of five PSIs for participation. Although for many of the analyses below this sample may be overpowered (i.e., able to detect statistically significant yet practically insignificant differences), the large sample was recruited because some of the behaviors reported on the surveys (e.g., suicide ideation and attempts) occur at very low frequencies. Issues related to power are discussed further in this chapter as well as in Chapter Five.

Figure 3.1 depicts the overall strategy of institution selection for the experimental pre-test. PSIs were invited to participate in the pre-test if they met the following criteria:

- They were already intending to participate in the ACHA-NCHA Spring 2007 data collection efforts;
- They were intending to survey students between February and May 2007;
- They were planning on using the online version of the ACHA-NCHA (the ACHA-NCHA is also available in paper version);
- They met eligibility requirements for ACHA-NCHA reference group inclusion (i.e., were planning on using random sampling or census sampling procedures in their survey efforts);
- They were large enough to invite a minimum of 1,000 students to participate in each sample (with expected returns of at least 250 students in each sample).

As PSIs self-selected to participate in the Spring 2007 data collection efforts beginning in late 2006, ACHA researchers invited schools that met the aforementioned criteria to

participate in the pre-test. As an incentive for participation, the ACHA provided the

following to each PSI that elected to participate:

1. The revised survey was administered free of charge. Each PSI was responsible for fees associated with implementing the original ACHA-NCHA among its students, but all fees associated with implementing the revised survey (i.e., contacting/recontacting students and downloading/processing data) were waived;

2. A free SPSS dataset of all data collected from the revised survey and a free frequency report for each variable on the revised survey by gender were generated for each PSI. Hence, participating institutions received datasets and data reports for both versions of the survey. Because the revised survey contained many new questions, institutions that participated in the pre-test received data about their students that were not collected on the original version of the ACHA-NCHA (e.g., numerous mental health variables, new impediments to academic performance such as homesickness and discrimination, information about stalking; prescription drug abuse);

3. A matched incentive program was offered in which the ACHA provided up to $1,000 per school to use towards incentives to entice students to complete the survey. A number of institutions (and particularly PSIs choosing to use the web-based ACHA-NCHA) utilize incentives (e.g., raffles for gift cards, iPods, parking passes, or airline tickets) to encourage students to participate in data collection efforts and hence increase response proportions during the implementation of the ACHA-NCHA on their campuses. If a PSI selecting to participate in the pre-test was already intending to provide incentives to its participating students in the original version of the survey, the ACHA matched the value of the incentives up to $1,000 in order to provide identical incentives to students participating in the revised version of the survey.

Once the minimum of 5 PSIs was reached, ACHA did not continue recruiting institutions;

ultimately, as shown in Figure 3.1, a convenience sample of 7 PSIs self-selected to

participate in the experimental pre-test. An additional institution did opt to use the

original version of the ACHA-NCHA to survey undergraduates and the revised version to

survey graduate students, but because survey strategies were not identical or controlled

across versions data collected from this PSI were not included in this study.

Characteristics of the 7 participating institutions are provided in Table 3.1. For

comparison purposes, characteristics of the 107 PSIs constituting the reference group are

provided as well (it should be noted that the 7 pre-test schools are included in the 107
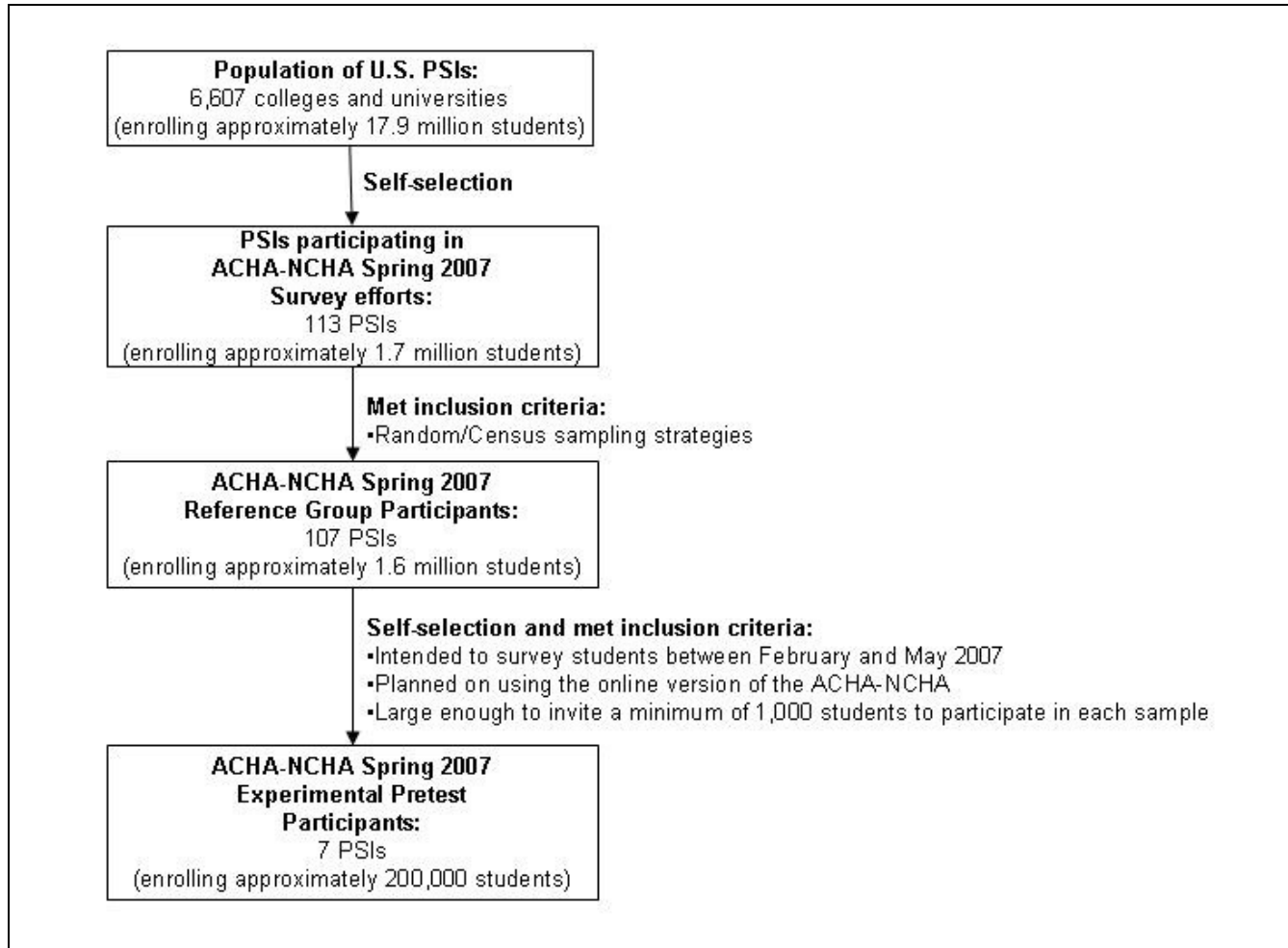
Figure 3.1. Selection strategy for PSIs to be included in the ACHA-NCHA survey experimental pre-test.

Table 3.1. Characteristics of PSIs participating in the Spring 2007 ACHA-NCHA experimental pre-test and Spring 2007 ACHA-NCHA reference group*.

Campus Characteristic	ACHA-NCHA Experimental Pre-test Total N=7 PSIs		ACHA-NCHA Reference Group Total N=107 PSIs	
	n	(%)	n	(%)
Type of Institution				
Public	6	(85.7)	69	(64.5)
Private	1	(14.2)	38	(35.5)
2-year	0	(0.0)	14	(13.1)
4-year	7	(100.0)	93	(86.9)
Location of Campus				
Northeast (CT, ME, MA, NH, NJ, NY, PA, RI, VT)	1	(14.2)	24	(22.4)
Midwest (IL, IN, IA, KS, MI, MN, MO, NE, ND, OH, SD, WI)	3	(42.9)	29	(27.1)
South (AL, AR, DE, DC, FL, GA, KY, LA, MD, MI, NC, OK, SC, TN, TX, VA, WV)	2	(28.6)	20	(18.7)
West (AK, AZ, CA, CO, HI, ID, MT, NV, NM, OR, UT, WA, WY)	1	(14.2)	32	(29.9)
Outside US	0	(0.0)	2	(1.9)
Campus Size				
< 2,500 students	0	(0.0)	16	(15.0)
2,500 – 4,999 students	0	(0.0)	10	(9.3)
5,000 – 9,999 student	0	(0.0)	20	(18.7)
10,000 – 19,999 students	1	(14.2)	25	(23.4)
20,000 students or more	6	(85.7)	36	(33.6)
Campus Setting				
Very large city (population over 500,000)	1	(14.2)	21	(19.6)
Large city (population 250,000-499,999)	2	(28.6)	14	(13.1)
Small city (population 50,000-249,999)	3	(42.9)	36	(33.6)
Large town (population 10,000 – 49,999)	1	(14.2)	26	(24.3)
Small town (population 2,500-9,999)	0	(0.0)	8	(7.5)
Rural community (population under 2,500)	0	(0.0)	2	(1.9)
Carnegie Classification				
Associates Colleges	0	(0.0)	14	(13.1)
Baccalaureate Colleges	0	(0.0)	14	(13.1)
Masters Colleges and Universities	0	(0.0)	35	(32.7)
Research Institutions	7	(100.0)	41	(38.3)
Special Focus Institutions	0	(0.0)	1	(0.9)
Miscellaneous/Not Classified	0	(0.0)	2	(1.9)
ACHA Membership Status				
Institutional Member	7	(100.0)	102	(95.3)
Nonmember	0	(0.0)	5	(4.7)
Religious Affiliation				
No	7	(100.0)	85	(79.4)
Yes	0	(0.0)	22	(20.6)
Postsecondary Minority Institution (US Dept. of Education)				
No	7	(100.0)	96	(89.7)
Yes	0	(0.0)	11	(10.3)

Note: PSIs participating in the ACHA-NCHA pre-test are a subset of PSIs participating in the Spring 2007 ACHA-NCHA reference group.

reference group schools as their students' responses on the original survey were included in the Spring 2007 reference group). In order to protect the identities of participating institutions, the ACHA only provides general information about each PSI.

Recruitment of student participants from participating institutions

As indicated above, a requirement for participation in the ACHA-NCHA experimental field pre-test was the use of generalizable sampling techniques of students. Hence, each PSI that self-selected to participate in the experimental pre-test of the ACHA-NCHA was responsible for providing the ACHA with a random sample or census list of its students' email addresses. The students on each list were then randomized to be invited to participate in either the original or revised survey, which ultimately created samples that were matched by institution for each version of the instrument. As is common practice for surveying efforts among college students, study participants were recruited to complete the online survey via the email address registered with their institutions. Online surveying is seen as a relevant means by which to collect health data (Daley et al., 2003; Pealer et al., 2001; Kypri et al., 2004) and is also very appropriate for measuring college students' experiences (Ouimet et al., 2004). No personal identification information (e.g., student identification number, full name, additional contact information) was linked to student email addresses, and email addresses were never linked to student responses. Each institution that participated in the spring 2007 ACHA-NCHA was required to provide the ACHA with Institutional Review Board (IRB) or administrative approval prior to survey implementation. Further descriptions of human subjects protection are described later in this chapter.

Students from each PSI were recruited to participate in both versions of the

survey using an email letter of invitation that was generated by the institution's contact

person, typically a student health representative, faculty member, or administrator.

Though letters differed between institutions, the letter was constant for both versions of

the survey at each PSI. All letters of invitation were reviewed by ACHA researchers to

ensure they included all relevant information (e.g., the length of time the survey would be

open, incentives provided) and as indicated above, all letters required IRB or

administrative approval from institutions before they were emailed to students. A

template letter of invitation, as provided in the ACHA-NCHA user's manual (ACHA,

2004b) is provided in Figure 3.2. The ACHA does not release specific letters of invitation

to protect the confidentiality of the institutions that participate in data collection efforts.

However, in general, all letters of invitation—regardless of institution—contained the

following information:

- The purpose of the study and its importance;
- Survey confidentiality and the means through which it was to be maintained;
- Benefits and risks;
- Informed consent procedures;
- Contact information for PSI representative for the ACHA-NCHA;
- Information about incentives and how they were to be awarded;
- Information about how long the survey will be available for students to complete; and
- A statement requiring that students must be 18 years of age or older to complete the survey.

Participating institutions were also required to provide the ACHA with a subject

line to be used for each email contact to students as well as for each reminder (i.e., non-

responder contact). Again, to protect the confidentiality of participating institutions, the

ACHA does not release the subject lines for the individual institutions that participated,

often because these subject lines contain the name of the PSI. It should be noted that the

ACHA discourages PSIs from using anything that looks like "spam" (e.g., numerous

Dear Student,

You have been randomly selected to participate in the National College Health Assessment Web survey (NCHA-Web) sponsored by the American College Health Association (ACHA). The NCHA-Web is a survey designed to assess student health behaviors in order to provide better services and support for *SCHOOL NAME* students. You may benefit by knowing that you have assisted in providing accurate information regarding health behaviors on your campus. The information will be used to develop
health programs for your campus.

The NCHA-Web is completed online via the Internet. You may scroll through the survey as you fill this out. We encourage you to complete the survey in one sitting, which typically takes about 20-30 minutes.

[Incentive information]

There may be some personal discomfort with the content of certain questions. For example, there are questions regarding illegal behaviors such as illegal substance use and sexual behavior. Your participation is completely voluntary and confidential. Your name or email addresses will never be associated with your responses. You may answer only some questions, or you may choose not to participate in the survey at all.

You have been assigned an ID number in order for the secure Internet server to manage your online survey input. This number is imbedded in your URL address. To ensure confidentiality, ID numbers and e-mail
addresses are destroyed before data are compiled and shared with your college or university.

Data transmission is encrypted and firewall securities are in place. After you submit the survey to secure server, you will receive a message thanking you for taking the NCHA-Web. The final survey responses will be housed at ACHA. Again, the version of the data set that is forwarded to your institution will not include personal identification such as e-mail addresses or ID numbers.

More directions follow as you link to the web site. By linking to the survey web site you are acknowledging that you are 18 years of age or older, and you are agreeing to participate in the NCHA-Web.

[Local contact information]

If you agree to participate in the ACHA NCHA-Web survey, click on the following Internet address to continue:

[URL HERE]

Thank you for you cooperation!
American College Health Association

Figure 3.2. Example letter of invitation template for the ACHA-NCHA web survey. Replicated with permission from the ACHA-NCHA User's Manual (ACHA, 2004b).

exclamation points, the phrase "win prizes," miscellaneous characters). A typical subject

line used is "*University X*: National College Health Assessment," though subject lines

vary greatly among participating institutions. Again, though each of the 7 participating

institutions differed in the subject lines used, the subject lines were kept constant between

versions of the survey at each school.

Once each participating PSI provided the list of student email addresses, required

text of the letter of invitation, and subject lines to the ACHA, the ACHA-NCHA Program

Office staff (including this researcher) was then responsible for programming each

school's individual survey and contacting student responders. The ACHA utilizes an

online survey program to assist with all aspects of the online surveying efforts—see

Appendix D, the ACHA-NCHA Frequency Asked Questions document for web

surveyors, and the ACHA-NCHA Users manual (ACHA, 2004b) for more information on

ACHA-NCHA survey practices. Each institution had a unique survey link—one for each

version of the survey—and links remained "live" for an approximate average of one

month. The letter of invitation from the PSI indicated the survey close date to its students.

Students from each PSI were surveyed between February and May 2007, although the

exact time period during which they were survey varied across institutions because of

factors such as spring break, the start and end of the spring semester, and timing of

midterms and final examinations.

A variety of strategies was used by each PSI to encourage student participation in

data collection efforts. Each PSI that participated in the ACHA-NCHA experimental pre-

test contacted non-responders at least one time (and as many as 3 times) to encourage

more students to complete the surveys. On average, students were given approximately

one month to complete the surveys, and non-responders were contacted within one or two weeks of the initial invitation. Contacting non-responders is a technique commonly utilized in the field of survey research, and it has been shown to significantly increase response proportions (e.g., de Vaus, 1995; Di Iorio, 2005). In addition to contacting non-responders, 6 of the 7 institutions that participated in the experimental pre-test also opted to provide an incentive for student participation. The ACHA will not release information on specific incentives provided, but common incentives used by participating institutions include raffles for gift certificates, cash prizes, and/or school paraphernalia (e.g., sweatshirts, t-shirts).

The number of students invited from each institution, as well as the final number of student respondents from each PSI is provided in Table 3.2. When samples from all institutions were combined, 71,023 students comprised the overall experimental pre-test sample. Of those, 533 total student email addresses bounced (less than one percent of the overall sample)—276 from the revised sample and 257 from the original sample. Accordingly, email invitations arrived to a total of 70,490 college students: 35,235 students in the revised survey sample and 35,255 in the original survey sample. A total of 6,216 students completed the revised version of the ACHA-NCHA and a total of 6,110 students completed the original version of the ACHA-NCHA. The overall response proportions for the revised and original versions were 17.6 percent and 17.3 percent, respectively, when excluding bounced email addresses from the eligible sample. The range of individual institution's response proportions was 7.4 – 34.6 percent for the revised survey (mean response proportion: 19.0 percent across PSIs) and 8.4 – 33.7 percent for the original survey. This response proportion, though low, is consistent with

Table 3.2. Student samples and response proportions for each of the 7 PSIs participating in the ACHA-NCHA experimental pre-test.

PSI ID	Total # of students in initial sample	# Students invited to participate in revised ACHA-NCHA (final sample excluding bounced emails)	# Completed revised ACHA-NCHA surveys	Revised ACHA-NCHA response proportion*	# Students invited to participate in original ACHA-NCHA (final sample excluding bounced emails)	# Completed original ACHA-NCHA surveys	Original ACHA-NCHA response proportion*
A	5,000	2,500 (2,500)	388	0.16	2,500 (2,500)	367	0.15
B	6,000	3,000 (2,928)	329	0.11	3,000 (2,907)	351	0.12
C	25,023	12,511 (12,511)	2,061	0.16	12,512 (12,512)	2,001	0.16
D	10,000	5,000 (4,947)	682	0.14	5,000 (4,970)	630	0.13
E	11,000	5,500 (5,429)	402	0.07	5,500 (5,427)	453	0.08
F	8,000	4,000 (3,970)	1,332	0.34	4,000 (3,973)	1,308	0.33
G	6,000	3,000 (2,950)	1,022	0.35	3,000 (2,966)	1,000	0.34
TOTAL	71,023	35,511 (35,235)	6,216	0.18	35,512 (35,255)	6,110	0.17

Note: Response proportion is calculated as the number of completed ACHA-NCHA surveys divided by the final eligible sample (i.e., the sample excluding bounced email addresses)

current literature in regards to online health surveys (Couper, 2007). Issues of non-response are discussed in more depth in Chapter Five. Characteristics of the final student samples for each version of the survey are provided in the next section of this chapter.

Characteristics of final student samples

As provided in Table 3.2, a total of 12,326 students participated in the ACHA-NCHA experimental pre-test study, 6,216 of whom completed the revised ACHA-NCHA and 6,110 of whom completed the original ACHA-NCHA. Because students were randomized to complete either the original or the revised survey instrument, and many factors were controlled in the study (e.g., PSIs used same subject lines, letters of invitation, and timing of survey administration across versions), it is not surprising that the final student samples across survey versions are very similar demographically. Table 3.3 provides an overview of the final demographics of the student samples from both the revised and original versions of the ACHA-NCHA. For comparison purposes, demographic information from all students who were included in the Spring 2007 ACHA-NCHA reference group is also included.

The only demographic characteristic on which the revised ACHA-NCHA sample and the original ACHA-NCHA sample differ significantly is the proportion of international students (revised: 9.5% international students, original: 6.3% international students). There are many differences between the experimental pre-test samples and the overall reference group sample, but interestingly, in many ways the pre-test samples are closer demographically to the population estimates particularly for gender breakdown, age distribution, and race/ethnicity for students attending 4-year PSIs as described earlier in this chapter (Kanpp et al., 2007; U.S. Census Bureau, 2006).

Table 3.3. Demographic characteristics of the experimental pre-test samples (revised and original) in comparison to the Spring 2007 ACHA-NCHA reference group sample.

Student Characteristic	Revised ACHA-NCHA Pre-test Sample Total N = 6,216		Original ACHA-NCHA Pre-test Sample Total N = 6,110		ACHA-NCHA Spring 2007 Reference Group Total N = 71,980	
	n	(valid %)	n	(valid %)	n	(valid %)
Gender						
Female	3,642	(59.6)	3,682	(61.0)	44,442	(64.1)
Male	2,469	(40.4)	2,352	(39.0)	24,932	(35.9)
Year in School						
First-year	1,216	(20.1)	1,164	(19.4)	17,437	(25.3)
Second-year	1,009	(16.6)	1,009	(16.8)	15,748	(22.8)
Third-year	1,360	(22.4)	1,330	(22.2)	15,086	(21.9)
Fourth-year	1,043	(17.2)	1,097	(18.3)	11,955	(17.3)
Fifth-year or more	326	(5.4)	284	(4.7)	3,704	(5.4)
Graduate/Professional	1,085	(17.9)	1,055	(17.6)	3,961	(5.7)
Other	25	(0.4)	51	(0.9)	1,088	(1.6)
Age						
18-23 years	4,894	(79.7)	4,849	(80.1)	59,882	(85.4)
24-34 years	1,050	(17.1)	995	(16.4)	7,908	(11.3)
35 years and above	199	(3.2)	210	(3.5)	2,303	(3.3)
Race/Ethnicity*						
White	4,350	(70.0)	4,318	(70.7)	52.678	(73.3)
Black	323	(6.0)	346	(5.7)	3,732	(5.2)
Hispanic/Latino(a)	370	(6.0)	393	(6.4)	5,832	(8.1)
Asian/Pacific Islander	999	(16.1)	975	(16.0)	7,922	(11.0)
American Indian/Alaskan Native	52	(0.8)	52	(0.9)	953	(1.3)
Other	183	(2.9)	228	(3.7)	2,955	(4.1)
International Student Status**						
No	5,503	(90.5)	5,632	(93.7)	65,359	(96.1)
Yes	580	(9.5)	378	(6.3)	2,704	(3.9)
Enrollment Status						
Full-time	5,686	(93.9)	5,661	(94.2)	65,521	(93.4)
Part-time	370	(6.1)	349	(5.8)	4,625	(6.6)
Residence						
Residence hall	1,946	(31.8)	1,860	(30.8)	26,326	(37.4)
Fraternity/Sorority house	119	(1.9)	128	(2.1)	868	(1.2)
Other on-campus housing	543	(8.9)	536	(8.9)	3,383	(4.8)
Off-campus	2,626	(43.0)	2,652	(43.9)	25,310	(35.9)
With parents	633	(10.4)	580	(9.6)	10,897	(15.5)
Other	244	(4.0)	285	(4.7)	3,631	(5.2)
Sexual Orientation						
Heterosexual	5,748	(94.4)	5,659	(94.1)	65,439	(93.6)
Gay/Lesbian	109	(1.8)	120	(2.0)	1,384	(2.0)
Bisexual	163	(2.7)	150	(2.5)	1,865	(2.7)
Unsure	71	(1.2)	83	(1.4)	1,202	(1.7)

***Note:** Percentages for race/ethnicity sum to more than 100% because students may select all that apply.
**Proportion of international students in the revised ACHA-NCHA pre-test sample was significantly different from the proportion in the original ACHA-NCHA pre-test sample (χ^2 = 43.648, df = 1, p<.001, Φ = .060)

Data Collection

The data analyzed in this study were collected using two versions of the ACHA-National College Health Assessment (see Appendices B.1 and B.2 both versions) from February to May 2007 by the ACHA. Both versions of the survey were web-based and took an estimated 20-30 minutes for students to complete (ACHA, 2007b). Each version of the survey was designed to collect information from college students on six content areas: (1) health, health education, and safety; (2) alcohol, tobacco, and other drug use; (3) sexual behavior; (4) mental and physical health; (5) impediments to academic performance; and (6) demographic characteristics. Formatting across surveys (e.g., font size, survey background, survey title) was as consistent as possible to ensure comparability across versions. As described in ACHA (2004b), earlier sections of this chapter, and in Appendix D, the ACHA utilized a specialized software application package to create the online instruments, and to collect and download data from students. Though each PSI collected data using the same surveys (i.e., the original and revised ACHA-NCHA), unique versions of both were created for each participating PSI to ensure the data were collected from only those students attending each institution. Further, unique URLs were generated for each sampled student to eliminate duplicate responses, provide the opportunity to contact non-responders, and to award incentives.

The focus of this dissertation research is to evaluate differences in student responses in the mental health indicators from both versions of the ACHA-NCHA, as well as to begin to evaluate evidence of preliminary reliability and validity of data collected from the revised mental health indicators. Subsequently, not every question on each version of they survey is of interest in this study. The following sections provide an

overview of the specific ACHA-NCHA questions from which data were analyzed for this study.

Revised ACHA-NCHA Survey Instrument

The revised ACHA-NCHA online survey instrument consists of 64 questions, many of which are in matrix form; subsequently, the survey actually collects 293 variables from student respondents. As described in Chapter One, the revised survey was developed as a result of feedback from the original ACHA-NCHA and changing college health priorities. Eight college health professionals, many of whom are considered experts in the field, and three ACHA staff members—including this researcher—formed the workgroup that revised the survey (see Appendix C for the list of revisions committee members). In consultation with the literature, theory, students from their respective institutions, and other college and mental health professionals, the workgroup designed the following questions to assess the mental health of U.S. college students:

- **NQ29:** Have you ever: (*Never, Not in the last 12 months, In the last 12 months, In the last 30 days, In the last 2 weeks*)
    - A. Felt things were hopeless
    - B. Felt overwhelmed by all you had to do
    - C. Felt exhausted (not from physical activity)
    - D. Felt very lonely
    - E. Felt very sad
    - F. Felt so depressed that it was difficult to function
    - G. Felt overwhelming anxiety
    - H. Felt overwhelming anger
    - I. Intentionally cut, burned, bruised, or otherwise injured yourself
    - J. Seriously considered suicide
    - K. Attempted suicide

- **NQ30:** Within the last 12 months, have you been diagnosed or treated by a professional for any of the following? (*No; Yes, diagnosed but not treated; Yes, treated with medication; Yes, treated with psychotherapy; Yes, treated with medication and psychotherapy; Yes, other treatment*)
    - A. Anorexia
    - B. Anxiety
    - C. Attention Deficit and Hyperactivity Disorder (ADHD)

    D. Bipolar Disorder
    E. Bulimia
    F. Depression
    G. Insomnia
    H. Other sleep disorder
    I. Obsessive Compulsive Disorder (OCD)
    J. Panic attacks
    K. Phobia
    L. Schizophrenia
    M. Substance abuse or addiction (alcohol or other drugs)
    N. Other addiction (e.g., gambling, internet, sexual)
    O. Other mental health condition

- **NQ31:** Have you ever been diagnosed (by a professional) with depression? (*No, Yes*)

- **NQ32:** Within the last 12 months, have any of the following been difficult to handle for you? (*No, Yes*)
    A. Academics
    B. Career related issue
    C. Death of a family member or friend
    D. Family problems
    E. Other relationships
    F. Finances
    G. Health problems of a family member or partner
    H. Personal appearance
    I. Personal health issue
    J. Other

- **NQ33:** Have you ever received psychological or mental health services from any of the following? (*No, Yes*)
    A. Counselor/Therapist/Psychologist
    B. Psychiatrist
    C. Other medical provider (e.g., physician, nurse practitioner)
    D. Minister/Priest/Rabbi/Other clergy

- **NQ34:** Have you ever received psychological or mental health services from your current college/university's Counseling or Health Service? (*No, Yes*)

- **NQ35:** If in the future you were having a personal problem that was really bothering you, would you consider seeking help from a mental health provider? (*No, Yes*)

- **NQ36:** Within the last 12 months, how would you rate the overall level of stress you have experienced? (*No stress, Less than average stress, Average stress, More than average stress, Tremendous stress*)

Altogether, the revised ACHA-NCHA contains 8 questions (4 matrix and 4 individual) that collect 44 variables worth of mental health information from students. The data collected from these 8 questions and corresponding 44 variables were used in analyses and evaluated in this dissertation. To the extent possible, revised items were compared to their counterparts on the original ACHA-NCHA. In addition, data collected from new items (i.e., items for which there are no comparable questions on the original survey) were evaluated for evidence of construct- and criterion-related validity.

Original ACHA-NCHA Survey Instrument

The original ACHA-NCHA began surveying college students in 2000 and stemmed largely from the CDC's National Collegiate Risk Behavior Survey (Douglas et al., 1997; CDC, 1997). The survey was initially only available in pen-and-paper format, but in Spring 2003 the ACHA began offering it in an online format as well (ACHA, 2007b). To date, more than 450,000 students from over 400 unique institutions have completed the survey. The original ACHA-NCHA online survey instrument consists of 72 questions, many of which are in matrix form; subsequently, the survey actually collects 273 variables from student respondents.

As the survey was first created in 1999—and before the emergence of much of the college student mental health crisis—there are a limited number of questions asking about student mental health concerns, which are provided below:

- **Q42:** Within the last school year how many times have you: (*Never, 1-2 times, 3-4 times, 5-6 times, 7-8 times, 9-10 times, 11 or more times*)
    - A. Felt things were hopeless
    - B. Felt overwhelmed by all you had to do
    - C. Felt exhausted (not from physical activity)
    - D. Felt very sad
    - E. Felt so depressed that it was difficult to function
    - F. Seriously considered attempting suicide

          G. Attempted suicide

- **Q43:** Have you ever been diagnosed with depression? (*Yes, No*)

- **Q44:** If yes: (*No, Yes*)
  - A. Have you been diagnosed with depression within the last school year?
  - B. Are you currently in therapy for depression?
  - C. Are you currently taking medication for depression?

- **Q47:** Within the last school year, have you had any of the following? (*No, Yes*)
  - B. Anorexia
  - C. Anxiety disorder
  - D. Bulimia
  - F. Depression

- **Q48:** Within the last school year, have you had any of the following? (*No, Yes*)
  - C. Substance abuse problem

Altogether, the original online ACHA-NCHA contains 5 questions that collect 16

variables worth of mental health information from students. The data collected from these

5 questions and corresponding 16 variables were compared to the data collected from the

revised ACHA-NCHA in this study when applicable.

Data Analyses and Evaluating Hypotheses

       All data analyses were conducted using Statistical Package for the Social Sciences

([SPSS], 2006) version 15.0.1. An overview of the procedures and analyses to be used for

the evaluation of each of the research questions and hypotheses are provided below.

Research questions 1 and 1A

       Because many of the revised mental health indicators were new or were

dramatically revised, and because data collected from the revised ACHA-NCHA survey

have not been explored in great depth previously, it was first of great interest to

determine what patterns of student responses emerged on each of these indicators. As

such, to evaluate research question 1, descriptive statistics (e.g., frequencies) were

generated for each of the mental health questions described earlier in this chapter. To evaluate research question 1A, which asks about how such endorsements of each item vary according to demographic characteristics, further frequency reports were generated that are stratified by such variables as gender, year in school, and race/ethnicity. As described in Chapter Two, certain mental disorders (e.g., anxiety, depression, eating disorders) are more common in female college students than male college students (NIMH, 2007b, 2007c) and others (e.g., substance abuse) are more common in male college students (Johnston et al., 2006; Presley et al., 1998). Therefore, the emergence of such patterns was expected in these data to provide further evidence of validity. It should be noted that patterns were not examined for statistical significance by demographic groups largely because of issues related to complex sampling; rather, these analyses were meant to merely explore the data without decisive conclusions by group.

Testing hypothesis 2

The literature suggests that even minor changes in survey questions create significant differences in respondent response patterns (e.g., Fowler, 1995; Fowler, 2004); therefore, it was hypothesized that changes in question formatting and response options between the original and revised ACHA-NCHA mental health indicators would result in significant differences in patterns of student responses on comparable items across versions of the survey. To evaluate this hypothesis, 2×2 contingency tables were created for each of the comparable survey items in which one variable in the table was the *version of the survey* (revised or original) and the other was the *pattern of student responses*. Data from each version of the survey were collapsed to form two response categories for comparison (e.g., "In the last 12 months, Not in the last 12 months" on the

revised survey vs. "In the last school year, Not in the last school year" on the original

survey). Table 3.4 provides an overview of the comparisons that were made across

versions of the survey in these analyses and how data were collapsed. It should be noted

that data were not artificially collapsed into variables with two response options—rather,

because of wording or response option changes each of the variables of interest for

comparison could only be directly compared when this procedure is used.

The issues related to complex sampling are not relevant for the purposes of these

analyses because students from each institution were randomized to receive either the

original or the revised survey. Hence, the variability that might be attributed to "school"

in these specific analyses is controlled through this randomization process.

A Pearson's chi-square (χ^2) for independent proportions statistic was calculated

for each comparison to determine if a relationship existed between version of the survey

and pattern of student responses; in other words, this test evaluated if any observed

differences in proportions between versions of the survey are statistically significant.

Fowler (2004) claims this type of analysis is appropriate to determine if changes in

survey questions produce significant differences in response patterns (see Chapter Two

for more information). Researchers suggest chi-square is more likely to establish

significance to the extent that (1) the relationship is strong, (2) the sample size is large,

and/or (3) the number of values of the two associated variables is large (Garson, n.d.). A

chi-square probability of .05 or less is commonly interpreted by social scientists as

justification for rejecting the null hypothesis that the row variable in the contingency

table is unrelated to the column variable (Garson).

Table 3.4. Comparisons made from data collected from mental health indicators across versions of the ACHA-NCHA.

Items from the original ACHA-NCHA used in comparisons	Items from the revised ACHA-NCHA used in comparisons	Data transformation procedures for comparisons
Q42: Within the last school year how many times have you: (*Never, 1-2 times, 3-4 times, 5-6 times, 7-8 times, 9-10 times, 11 or more times*) A. Felt things were hopeless B. Felt overwhelmed by all you had to do C. Felt exhausted (not from physical activity) D. Felt very sad E. Felt so depressed that it was difficult to function F. Seriously considered attempting suicide G. Attempted suicide	**NQ29:** Have you ever: (*Never, Not in the last 12 months, In the last 12 months, In the last 30 days, In the last 2 weeks*) A. Felt things were hopeless B. Felt overwhelmed by all you had to do C. Felt exhausted (not from physical activity) E. Felt very sad F. Felt so depressed that it was difficult to function J. Seriously considered suicide K. Attempted suicide	• Original response options were collapsed to "Never within the last school year" (*Never*) and "Once or more during the last school year" (*All other response options*) • Revised response options were collapsed to "Not in the last 12 months" (i.e., *Never* and *Not in the last 12* months) and "In the last 12 months" (*All other response options*)
Q43: Have you ever been diagnosed with depression? (*Yes, No*)	**NQ31:** Have you ever been diagnosed (by a professional) with depression? (*No, Yes*)	• Proportions of students who respond *Yes* and *No* to these items were directly compared.

Q47: Within the last school year, have you had any of the following? (*No, Yes*)
  B. Anorexia
  C. Anxiety disorder
  D. Bulimia
  F. Depression
 	
Q48: Within the last school year, have you had any of the following? (*No, Yes*)
  C. Substance abuse problem
	NQ30: Within the last 12 months, have you been diagnosed or treated by a professional for any of the following?
	(*No; Yes, diagnosed but not treated; Yes, treated with medication; Yes, treated with psychotherapy; Yes, treated with medication and psychotherapy; Yes, other treatment*)
	  A. Anorexia
	  B. Anxiety
	  D. Bulimia
	  F. Depression
	  M. Substance abuse or addiction (alcohol or other drugs)
		• Original response options remained as *No* and *Yes*
		• Revised response options were collapsed to "No" (*No*) and "Yes" (*All yes responses combined*)

Generally speaking, the chi-square test for independent proportions is a non-parametric statistical test used to make inferences about categorical data. When patterns observed in the data differ from patterns expected in the data (i.e., the patterns that would exist if no association between the two variables was present), the observed chi-square statistic will exceed the critical chi-square statistic (at α = .05 and df=1, $\chi^2_{critical}$ = 3.841). Assumptions for the chi-square test for independence are as follows (Garson, n.d.), all of which are expected to be met by the data collected during the ACHA-NCHA experimental field pre-test:

- Data collected are a random sample of a larger population;
- Data are from a sufficiently large sample size. Although there is no accepted cutoff, it is suggested that a minimum of between 20 and 50 data records should be collected to perform the test;
- Adequate cell sizes are present. This commonly equates to cell sizes of 5 or more in all cells of a 2×2 contingency table;
- Observations are independent; in other words, each subject contributes data to only one cell;
- Observations must have the same underlying distribution;
- The hypothesized distribution is specified in advance, such that the number of observations that are expected to appear in each cell in the table can be calculated without reference to the observed values. Normally this expected value is the cross-product of the row and column marginals divided by the sample size;
- Non-directional hypotheses are assumed;
- Observations are in the form of categorical data; and
- Deviations (observed minus expected values) are normally distributed.

Because the overall sample was very large (i.e., more than 6,000 respondents per version of the survey), it was unnecessary to use the Yates correction often used in the analysis of 2×2 contingency tables (Pagano & Gauvreau, 2000; Field, 2005). The standard level of Type I error (α = .05) was used in these analyses. However, because of the very large sample sizes, there was sufficient power to detect even very small differences in proportions (Garson, n.d.); as such, Phi (Φ) coefficients were calculated to

determine the extent of the association between version of the survey and pattern of student responses (Field, 2005). Standard estimates of small (0.1), medium (0.3), and large (0.5) effects were used to determine overall level of association for the Phi coefficient (Miles & Gilbert, 2005). For the purposes of these analyses, even small effects ($\Phi = 0.1$ or higher) were of interest between versions of the survey. However, because Phi values lower than 0.1 were of little practical significance, only those comparisons that resulted in Phi estimates of 0.1 or greater were used to generate support for Hypothesis 2.

Testing hypotheses 3A and 3B

The third research question of interest in this study was "To what extent are the mental health indicators on the revised ACHA-NCHA reliable?" In general, reliability refers to the consistency and stability of measurement and is impacted by random measurement error. See Chapter Two for a more comprehensive overview of reliability and validity in survey research. It was hypothesized that (1) measurement scales on the revised ACHA-NCHA will demonstrate sufficient levels of internal consistency in the pre-test sample, and (2) reliability will not increase if any indicators are removed from measurement scales on the revised ACHA-NCHA in the pre-test sample.

To test these hypotheses, the internal consistency and other psychometric properties of the negative affect scale were evaluated. The negative affect scale on the revised ACHA-NCHA, which is presented below, consists of 8 indicator variables:

- **NQ29:** Have you ever: (*Never, Not in the last 12 months, In the last 12 months, In the last 30 days, In the last 2 weeks*)
    - Felt things were hopeless
    - Felt overwhelmed by all you had to do
    - Felt exhausted (not from physical activity)
    - Felt very lonely
    - Felt very sad
    - Felt so depressed that it was difficult to function

- Felt overwhelming anxiety
- Felt overwhelming anger

First, in order to determine the dimensionality of the negative affect scale on the revised ACHA-NCHA (i.e., if the scale is uni- or multidimensional), data were analyzed using exploratory factor analysis. Exploratory factor analysis is often used to evaluate psychometric properties of latent variables (i.e., variables that cannot be directly measured) and it is regularly used to evaluate relationships among data that are suspected to driven by the same underlying variable (Field, 2005). In this case, each of the eight indicator variables on the scale above was suspected to be driven by the underlying construct of a negative affect. An exploratory factor analysis allows researchers to explore data to discover clusters of large correlation coefficients between subsets of variables—these subsets of highly correlated variables are commonly referred to as "factors".

Specifically, the revised ACHA-NCHA negative affect scale was analyzed using a principal components analysis (PCA) with oblique rotation. PCA, a method used to reduce data, summarize data, and explain maximal variance in variable sets, is a very common statistical technique in the social sciences (Smith, 2002; Field, 2005). The basic premise of PCA is that it linearly transforms an original set of variables into a substantially smaller set of variables that represents most of the information in the original set (Dunteman, 1989). Oblique rotation allows for two or more components, or dimensions, to correlate with one another. If more than one component, or dimension, were to have emerged with this scale, it would have beed expected that they would correlate with one another. To determine the number of scale dimensions, the commonly accepted practice of examining a combination of evidence from the scree plot—

specifically the point of inflexion—and from eigenvalues greater than one (i.e., those

components that explain as much variance in the data as one of the standardized

variables) was used. The assumptions of PCA are as follows (Field, 2005):

- All variables included in the PCA are measured on an interval or ratio scale;
- The relationship between all observed variables is linear; and
- Each pair of observed variables displays a bivariate normal distribution.

Although the response options for the indicator variables on the revised negative affect

scale are more ordinal in nature than continuous, which violates the first assumption

listed above, PCA will likely still provide valuable and relevant information about the

relationships between these variables (Korhonen & Siljamaki, 1998).

Once the number of scale dimensions had been determined through the use of

PCA (only one dimension emerged, see Chapter Four for details), the reliability of the

scale was assessed. As described in Chapter Two, internal consistency reliability

estimates are perhaps the most commonly used in health survey research (Di Iorio, 2005).

These approaches estimate reliability by assessing correlations and relationships among

items within a scale. The reliability coefficient is then assessed by calculating the

coefficient alpha (Cronbach, 1951), a process through which each item is correlated with

each other item on a scale and alpha is computed by taking the average of individual

item-to-item correlations and adjusting for the number of items (Di Iorio). Higher

coefficient alphas are associated with scales that have more items and higher interitem

correlations, and it has been suggested that an alpha coefficient of .70 or greater

demonstrates sufficient reliability (Nunnally & Bernstein, 1994, as cited in Di Iorio,

2005) of a survey scale. In addition to examining the coefficient alpha for each emergent

scale, the output was further examined to determine if alpha would increase if any of the

indicators are removed from the scale to evaluate Hypothesis 3B.

In summary, the basic steps that were used to evaluate the third research question in this study were as follows:

- A PCA with oblique—specifically, direct oblimin—rotation was used to determine the underlying dimensionality of the scale;
- The number of dimensions of the scale was determined through a combination of the examination of (1) the point of inflexion on the scree plot, and (2) the number of factors generated with an eigenvalue over one (Kaiser, 1960);
- A variable was said to be an indicator of the underlying scale dimension if it loaded onto the component with a coefficient of 0.4 or higher, a standard value in the field of PCA (Dunteman, 1989);
- Each subscale (if more than one emerges) was assessed using an internal consistency reliability analysis. Coefficient alpha was computed for each of the emergent subscales and subscales with alpha coefficients of .70 or higher was said to demonstrate sufficient evidence of reliability; and
- The extent to which each indicator variable contributes to the reliability of the subscale(s) was analyzed by examining whether coefficient alpha would increase if the indicator is removed from the scale.

As discussed in Chapter Two, reliability is a necessary but insufficient condition for validity. Hence, demonstrating the internal consistency of the negative affect scale was imperative if any claims about the validity of its measurement are to be made. The next research question and associated hypotheses refer to establishing evidence of validity in data collected from the negative affect scale as well as numerous other revised ACHA-NCHA mental health indicators.

Testing hypotheses 4A and 4B

The term validity refers to the legitimacy of an instrument—that is, does it measure what it purports to measure? More specifically, validity is the degree of support obtained for the interpretations of scores on an instrument when the instrument is used for its intended purpose (Di Iorio, 2005). The fourth research question of interest in this study examined the extent to which data collected from the revised ACHA-NCHA mental health indicators are valid; hence, strategies were utilized to begin to assess how well

these data appeared to (1) demonstrate evidence of criterion-related validity, and (2) demonstrate evidence of construct-related validity. A comprehensive overview of each of these types of validity is provided in Chapter Two; in addition, a brief summary of each is provided below with the description of the methodology that was used.

Criterion-related validity refers to the extent to which the data collected from a survey agrees with some "true" value or "gold standard" of measurement of a construct. In the field of health survey research, this may equate to medical records or directly observable health data. For example, to determine evidence of criterion-related validity, self-reports of depression may be correlated with diagnoses of depression in medical files. In the case of the revised ACHA-NCHA, there are no such data to which self-report indicators may be compared, particularly because no student identifiers have been linked to data and the ACHA cannot release which PSIs participated in the experimental pre-test to protect the confidentiality of the institutions. Therefore, a somewhat different yet commonly utilized strategy was utilized to evaluate evidence of criterion-related validity in these data: student responses from the revised ACHA-NCHA mental health indicators were compared to prevalence estimates collected from the nationally generalizable National Co-Morbidity Survey Replication ([NCS-R], Kessler et al., 2005) among people of relatively similar demographics. The literature suggests comparisons of data collected from new surveys to data collected from previously validated instruments is an appropriate method to begin to establish evidence of validity (DeVellis, 2003), though certainly these comparisons have a variety of limitations when samples, survey methodology, and survey questions are not directly comparable, as was the case in this study (see Chapter Five for more information).

The baseline NCS (Kessler, 1994) was conducted in 1990-1992 as a national survey effort to evaluate the mental health status of U.S. residents ages 15-54. This survey was a fully structured interview that was administered face-to-face using paper and pencil interviewing with trained lay interviewers. Data from this study are considered to be nationally generalizable and are among the most commonly cited in the literature in the field of mental health. The NCS-R Survey (Kessler et al., 2005) was conducted in 2001-2002 as a national survey of people ages 18 years and over. Like the baseline NCS, the NCS-R was a fully structured interview administered by trained lay interviewers. This study aimed to both replicate some aspects of the baseline NCS for trending purposes and to expand on the data available on the status of mental health in the United States. Twelve-month prevalence estimates are available by gender and by age for a variety of mental health conditions, including anxiety disorders, mood disorders, eating disorders, ADHD, and substance abuse disorders. These 12-month prevalence estimates were compared to data collected from the ACHA in the items below:

- **NQ30:** Within the last 12 months, have you been diagnosed or treated by a professional for any of the following? (*No; Yes, diagnosed but not treated; Yes, treated with medication; Yes, treated with psychotherapy; Yes, treated with medication and psychotherapy; Yes, other treatment*)
  - A. Anorexia
  - B. Anxiety
  - C. Attention Deficit and Hyperactivity Disorder (ADHD)
  - D. Bipolar Disorder
  - E. Bulimia
  - F. Depression
  - I. Obsessive Compulsive Disorder (OCD)
  - J. Panic attacks
  - K. Phobia
  - M. Substance abuse or addiction (alcohol or other drugs)

In order to generate the 12-month prevalence estimates for each of the conditions in NQ30, all "yes" responses were collapsed prior to comparison with prevalence estimates from the NCS-R. Prevalence estimates were compared by gender and by age group as

provided (ages 18-29 and 30-44). Though populations of interest are not identical (i.e., not all NCS-R participants ages 18-44 are enrolled in PSIs as they are in the ACHA-NCHA), which is certainly a limitation of this methodology, comparisons were useful to begin to evaluate evidence of criterion-related validity of the revised ACHA-NCHA data.

Criterion validity was, for many years, the mainstay of validity assessment (Di Iorio, 2005). However, in many cases, researchers realized there were often no suitable criteria for comparison, a phenomena that was particularly salient for the measurement of latent psychological variables such as negative affect, the measurement scale for which is described above. To address this problem, researchers developed strategies to evaluate construct validity. Because constructs rarely exist in isolation, survey methodologists suggest that if an instrument measures what it claims to measure, data from the instrument would exhibit patterns that would be expected according to theory (Di Iorio). When data confirm hypothesized relationships, they demonstrate evidence of construct-related validity (Aday & Cornelius, 2006).

Construct validity is comprised of both discriminant and convergent validities, evidence of which is frequently examined together and involves studying the strengths or patterns among data collected from a variety of variables (Spector, 1992). Convergent validity refers to a strong correlation between different measures of the same construct. Discriminant validity suggests measures of different constructs should relate only modestly with one another (Spector).

Typically, evidence of construct-related validity is demonstrated through the use of statistical models and analyses. Perhaps no statistical technique is as commonly used to measure associations and model relationships among variables as regression in the

social sciences. Hence, to evaluate evidence of construct-related validity of the data collected from the revised ACHA-NCHA, these data were fit to both multiple linear and logistic regression models. Multiple linear regression uses the method of least squares to explore the relationship between many predictor (independent) variables and one continuous outcome (dependent) variable. Linear regression attempts to model the relationship between variables by fitting a linear equation to observed data. Logistic regression, on the other hand, is used when outcome variables are dichotomous; this technique expresses the multiple linear equation in logarithmic terms and thus overcomes the problem of violating the assumption of linearity that would occur if linear regression to predict a dichotomous variable.

The assumptions of linear regression and logistic regression are provided in Table 3.5 below. As is evidenced in the table, unlike linear regression, logistic regression does not assume linearity of the relationship between the independent and dependent variables, normally distributed variables, or homoscedasticity.

Table 3.5. Assumptions of linear and logistic regression statistical techniques.

Linear Regression	Logistic Regression
• Outcome variable is continuous	• Outcome variable is categorical
• Observations are independent	• Observations are independent
• Independent variables are continuous or categorical	• Independent variables are continuous or categorical
• Independent variables are linearly related to the dependent variable	• Independent variables are linearly related to the logit of the dependent
• No multicollinearity	• No multicollinearity
• Errors are independent	• No outliers
• Errors are normally distributed	
• Homoscedasticity	

To evaluate evidence of construct-related validity, two linear regression models were evaluated in this study. Each utilized forced entry ordinary least squares (OLS) methods in order to examine relationships between the predictor variables and outcome variables based on theory. Though each model is examining only simple relationships, multiple regression was used because it provides the opportunity to control for blocking variables such as age, gender, year in school, and race/ethnicity. The two linear models that were tested are provided below:

- *Model A: Predicting negative affect from level of stress.* This model provided the opportunity to examine the ability of the independent variables below to explain the variance in self-reported negative affect and had the capacity to demonstrate evidence of construct-related validity because, according to theory and previous literature, stress predicts negative affect (e.g., Ross et al., 1999; Cook, 2007).

  o **Dependent variable:** Recency of negative affect, a continuous variable that was created by generating a factor score from the negative affect scale
  o **Predictors**
    ▪ Block 0: No predictors
    ▪ Block 1: Demographic characteristics for which to control (*Model A1*)
      - Age (continuous)
      - Year in school
      - Gender
      - Race/ethnicity
    ▪ Block 2: Level of stress (*Model A2→Included blocks 1 and 2 simultaneously*)
      - Student responses from NQ36 were entered into the model

- *Model B: Predicting level of stress from experience of a variety of difficult life events.* This model provided the opportunity to examine the ability of the five most commonly experienced difficulties in the last 12 months to explain the variance in self-reported stress levels and had the capacity to demonstrate

evidence of construct-related validity because, according to theory and previous

literature, the experience of difficult life events is theoretically linked to stress

levels (e.g., Ross et al., 1999).

- o **Dependent variable:** Self-reported level of stress, a 5 point continuous variable (NQ36)
- o **Predictors**
  - Block 0: No predictors
  - Block 1: Demographic characteristics for which to control (*Model B1*)
    - Age
    - Year in school
    - Gender
    - Race/ethnicity
  - Block 2: Stressful life events (*Model B2→Included blocks 1 and 2 simultaneously*)
    - Student responses from the 5 most commonly endorsed indicators from NQ32 were entered as a series of independent variables in this block.

Through the examination of the change in R^2 values (i.e., the amount of variance

in the outcome variable explained by the predictors of interest) from block to block, these

models allowed for the examination of (1) how various demographic characteristics

influence the outcomes of interest, and (2) how the theoretically linked constructs of

stress, stressful life events, and negative life events interrelate when controlling for the

aforementioned demographic traits. Though significant relationships (i.e., B coefficients)

were expected, thereby providing evidence of convergent validity, it was not expected

that all of the variance in the dependent variables would be explained by the predictors

specified in the models. Hence, because R^2 was not expected to equal 1, these models

possessed the capacity to provide further evidence of discriminant validity. This

simultaneous combination of discriminant and convergent validity evidence contributes

to evidence of construct-related validity in each of these mental health indicators.

In addition to the two linear regression models described above, the logistic regression model presented next was examined to evaluate evidence of construct-related validity. As was the case above, it utilized forced entry methodology in order to examine relationships between the predictor variables and outcome variable based on theory.

- *Model C: Predicting diagnosis with or treatment for depression in the last 12 months from negative affect and stress level*. This model provided the opportunity to examine the ability of the independent variables below to predict whether or not a student had been diagnosed with or treated for depression in the last year and had the capacity to demonstrate evidence of construct-related validity because, according to theory and previous literature, negative affect and stress have both been theoretically linked to depression in college students (NMHA, 2007b, Grace, 1997).

  - **Dependent variable:** Diagnosis (by a professional) with depression in the last 12 months (NQ31)
  - **Predictors**
    - Block 0: No predictors
    - Block 1: Demographic characteristics for which to control (*Model C1*)
      - Age
      - Gender
      - Year in school
      - Race/ethnicity
    - Block 2: Level of stress and recency of negative affect (*Model C2→Included blocks 1 and 2 simultaneously*)
      - Student responses from level of stress (NQ36)
      - Recency of negative affect, a continuous variable that was created by generating a factor score from the negative affect scale

As is evidenced above, this logistic regression model controlled for the blocking variables of gender, age, year in school, and race/ethnicity in the data sample. This is because mental health status has been shown to vary according to these demographic

characteristics; thus, in order to demonstrate the theoretical relationships between the constructs of interest, it was important to control for these variables. Statistics of interest in this analysis included b values, Exp(b) values (i.e., indicators of the change in odds of the event occurring resulting from a one unit change in the predictor of interest), and change in log likelihood values between models (when -2LL decreases between models, it indicates newly incorporated independent variables are improving predictive abilities and better overall model fit). Again, the combination of information collected was expected to demonstrate evidence of convergent and discriminant validity which could ultimately provide confirmation of construct-related validity for each of these indicators.

Testing Hypothesis 5

The final research question of this study, and perhaps that which will be of the greatest interest to many in the field of college student mental health, asked "how do mental health indicators on the revised ACHA-NCHA differ from those on the original ACHA-NCHA in terms of their reliability and validity?" The purpose of revising the ACHA-NCHA was to create a college health survey that reduces measurement error to every extent possible. Further, because such strenuous steps were taken to create a survey that was more comprehensive, clear, and accurate than the original, it was expected that data collected from the revised ACHA-NCHA would be more accurate and consistent than those collected from the original survey instrument. Therefore, it was hypothesized that mental health indicators on the revised ACHA-NCHA will demonstrate evidence of higher levels of reliability and validity than those on the original ACHA-NCHA.

In order to test this hypothesis, many of the procedures described above were further conducted using the original ACHA-NCHA mental health data, and comparisons

were made across versions of the survey. The evaluation of this research question focused

primarily on (1) differences in reliability of the negative affect scale across versions of

the survey, and (2) evidence of criterion-related validity of data collected on the

prevalence of mental health conditions. Coefficient alphas for the negative affect scale

were compared from the original to the revised version of the survey, and prevalence

estimates of anorexia, anxiety, bulimia, depression, and substance abuse disorders were

also compared to one another and to the NCS-R prevalence estimates to determine which

of the two ACHA-NCHA versions are more similar to those prevalence estimates.

Human Subjects Procedures

As this study is primarily a secondary data analysis of data collected by the

American College Health Association, human subjects concerns are minimal for the

purposes of this dissertation. Data are publicly available upon request to the ACHA. No

personal identifiers are included, and only general information in regards to which

universities were surveyed (e.g., Carnegie classification, general region) are available. As

such, no individual student nor participating institution can be identified through these

data. Hence, this research was deemed "exempt" by the University of Maryland

Institutional Review Board (IRB). A copy of the IRB approval letter for this research is

included as Appendix E.

It should further be noted that the ACHA requires IRB approval from each

institution that participates in the ACHA-NCHA prior to the commencement of survey

activities. The ACHA keeps records, which are not available to the public, of each

institution's IRB approval for its ACHA-NCHA survey efforts. Additionally, all ACHA

researchers have been trained on human subjects' protection through both the National

Institutes of Health and the Collaborative Institutional Training Initiative (CITI)

programs. Moreover, each PSI that participated in the online survey effort was required

to provide the ACHA with an online consent form that had been approved by its IRB.

Students were emailed these online consent forms/letters of invitation to participate in the

survey and were notified that they were giving their consent to participate by clicking on

the survey link provided in the email letter of invitation. Accordingly, there were multiple

strategies and processes in place to ensure the confidentiality of student data and

institutional privacy are maintained, and that the rights of human subjects are both

preserved and protected.

CHAPTER IV: RESULTS

Introduction

The aim of this study was to evaluate psychometric characteristics of the revised and newly added mental health indicators on the revised ACHA-NCHA survey. To accomplish these aims, this study (1) evaluated how changing the wording of questions designed to measure the same construct changed patterns of student responses across versions of the survey, and (2) examined evidence of internal consistency reliability and construct- and criterion-related validity of the data collected from the revised ACHA-NCHA. Finally, data collected from the revised and original ACHA-NCHA mental health indicators were compared in terms of their reliability and validity. To accomplish these aims, data collected from the original (n = 6,120) and the revised (n = 6,216) ACHA-NCHA online survey were analyzed using a variety of strategies. Seven U.S. PSIs participated in data collected efforts, and as described in depth in Chapter Three, students from each institution were randomized to participate in either the original or revised survey under controlled conditions.

The final samples were nearly identical demographically, and data from each were used in the analyses that follow. Before any results are presented, it should first be noted that no significant problems with item non-response appeared on either survey. The average non-response for mental health indicators on the revised ACHA-NCHA was 1.5 percent (minimum: 1.1 percent, maximum: 2.7 percent); the average non-response for mental health indicators on the original ACHA-NCHA was 1.1 percent (minimum: 0.9 percent, maximum: 1.5 percent). The greatest percent of non-response on the revised

ACHA-NCHA was observed for the item asking students if they had been diagnosed or treated in the last 12 months with an "other" mental health condition. The findings of this study are presented below and are organized by each research question and its associated hypotheses.

Research Questions 1 and 1A

Because many of the survey items asking about student mental health were new and had never been used to collect data, it was first of interest to determine general patterns of student responses to each of the 8 questions (4 individual and 4 matrix) and 44 associated variables measuring students' mental health status. Tables 4.1, 4.2, and 4.3 present frequency data by gender and in total for each of the mental health variables. Table 4.1 presents student responses to each of the items on the negative affect scale, the psychometric properties of which are evaluated later in this chapter. Table 4.2 presents information in regards to the prevalence of a variety of mental health conditions, stressful life events, and help-seeking behaviors. Table 4.3 summarizes deliberate self-harm as well as suicidality in the pre-test sample according to a variety of demographic characteristics, including gender, race/ethnicity, and year in school.

As is provided in Table 4.1, students most frequently reported feeling overwhelmed (n = 2,793; 45.5 percent) and exhausted—not from physical activity (n = 2,601, 42.4 percent) in the last 2 weeks. Fewer students reported feeling overwhelming anger (n = 596; 9.8 percent) or that they were so depressed it was difficult to function (n = 520; 8.5 percent) in the last 2 weeks compared to other indicators of negative affect. Female students reported experiencing each of these feelings with greater recency than

Table 4.1. Student responses to negative affect scale items by gender (*Stem*: "Have you ever:").

	Females		Males		Total	
	n	(valid %)	n	(valid %)	n	(valid %)
Felt things were hopeless						
• Never	994	(27.6)	1,014	(41.3)	2.028	(33.1)
• Not in the last 12 months	771	(21.4)	475	(19.3)	1,259	(20.5)
• In the last 12 months	895	(24.8)	478	(19.5)	1,392	(22.7)
• In the last 30 days	374	(10.4)	192	(7.8)	567	(9.3)
• In the last 2 weeks	572	(15.9)	297	(12.1)	881	(14.4)
Felt overwhelmed						
• Never	151	(4.2)	361	(14.7)	520	(8.5)
• Not in the last 12 months	103	(2.8)	171	(7.0)	275	(4.5)
• In the last 12 months	785	(21.7)	614	(25.0)	1,409	(22.9)
• In the last 30 days	695	(19.2)	441	(17.9)	1,148	(18.7)
• In the last 2 weeks	1,891	(52.2)	871	(35.4)	2,793	(45.5)
Felt exhausted						
• Never	255	(7.0)	436	(17.8)	702	(11.4)
• Not in the last 12 months	190	(5.2)	198	(8.1)	394	(6.4)
• In the last 12 months	691	(19.1)	526	(21.4)	1,227	(20.0)
• In the last 30 days	724	(20.0)	485	(19.8)	1,217	(19.8)
• In the last 2 weeks	1,764	(48.7)	810	(33.0)	2,601	(42.4)
Felt very lonely						
• Never	543	(15.0)	624	(25.4)	1,183	(19.3)
• Not in the last 12 months	681	(18.8)	476	(19.3)	1,165	(19.0)
• In the last 12 months	952	(26.3)	591	(24.0)	1,558	(25.4)
• In the last 30 days	574	(15.9)	304	(12.4)	887	(14.5)
• In the last 2 weeks	866	(23.9)	465	(18.9)	1,345	(21.9)
Felt very sad						
• Never	453	(12.5)	620	(25.2)	1,087	(17.7)
• Not in the last 12 months	572	(15.8)	464	(18.9)	1,047	(17.1)
• In the last 12 months	1,052	(29.1)	690	(28.1)	1,753	(28.6)
• In the last 30 days	610	(16.9)	285	(11.6)	904	(14.7)
• In the last 2 weeks	929	(25.7)	399	(16.2)	1,344	(21.9)
Felt so depressed it was difficult to function						
• Never	1,478	(40.8)	1,290	(52.6)	2,794	(45.5)
• Not in the last 12 months	860	(23.7)	525	(21.4)	1,398	(22.8)
• In the last 12 months	668	(18.4)	345	(14.1)	1,024	(16.7)
• In the last 30 days	273	(7.5)	125	(5.1)	402	(6.5)
• In the last 2 weeks	344	(9.5)	168	(6.8)	520	(8.5)
Felt overwhelming anxiety						
• Never	1,015	(28.0)	1,056	(43.1)	2,093	(34.1)
• Not in the last 12 months	585	(16.2)	427	(17.4)	1,020	(16.6)
• In the last 12 months	889	(24.6)	487	(19.9)	1,388	(22.6)
• In the last 30 days	502	(13.9)	229	(9.3)	740	(12.1)
• In the last 2 weeks	628	(17.4)	253	(10.3)	892	(14.5)
Felt overwhelming anger						
• Never	1,305	(36.3)	1,008	(41.1)	2,334	(38.2)
• Not in the last 12 months	766	(21.3)	539	(22.0)	1,315	(21.5)
• In the last 12 months	787	(21.9)	485	(19.8)	1,290	(21.1)
• In the last 30 days	372	(10.3)	195	(8.0)	575	(9.4)
• In the last 2 weeks	369	(10.3)	223	(9.1)	596	(9.8)

male students. When collapsing response categories to determine the frequency with

which each of these indicators was experienced in the last 12 months (i.e., combining

experienced *In the last 12 months; In the last 30 days,* and *In the last 2 weeks*), the

frequencies with which each was reported on the revised survey was, in order from most

frequently experienced through least frequently experienced (this order remains

unchanged regardless of gender):

- Felt overwhelmed (n = 5,350; 87.1 percent)
- Felt exhausted—not from physical activity (n = 5,045; 82.2 percent)
- Felt very sad (n = 4,001; 65.2 percent)
- Felt very lonely (n = 3,790; 61.7 percent)
- Felt overwhelming anxiety (n = 3,020; 61.7 percent)
- Felt things were hopeless (n = 2,840; 46.4 percent)
- Felt overwhelming anger (n = 2,461; 40.3 percent)
- Felt so depressed it was difficult to function (n = 1,946; 31.7 percent)

Table 4.2 provides an overview of the prevalence estimates (i.e., diagnosis or

treatment within the last 12 months) of a variety of mental health conditions in this

sample of more than 6,000 students. The top five mental health conditions reported were:

Depression (9.3 percent); Anxiety (8.5 percent); Panic attacks (4.0 percent); Insomnia

(3.2 percent); and Attention Deficit and Hyperactivity Disorder ([ADHD] 2.9 percent).

Women reported each of these five conditions with greater frequency than men with the

exception of ADHD; in fact, the prevalence estimate of having been diagnosed or treated

with depression and anxiety within the last 12 months was more than double for women

than it was for men (Depression: 11.7 percent and 5.7 percent, respectively; Anxiety:

10.7 percent and 5.0 percent, respectively) in this sample. Further, approximately 14.9

percent of this sample reported having ever been diagnosed (by a professional) with

depression. Again, this lifetime prevalence estimate was higher for women (18.2 percent)

than for men (10.0 percent) in this sample. In addition to ADHD, men reported bipolar

Table 4.2. Proportion of "yes" responses to select mental health indicators on the revised ACHA-NCHA by gender.

	Females (n=3,642)		Males (n=2,469)		Total (n=6,216)	
	n	(valid %)	n	(valid %)	n	(valid %)
Within the last 12 months, have you been diagnosed or treated by a professional with any of the following?*						
• Anorexia	37	(1.0)	17	(0.7)	57	(0.9)
• Anxiety	387	(10.7)	123	(5.0)	521	(8.5)
• Attention Deficit Hyperactivity Disorder	78	(2.2)	92	(3.7)	176	(2.9)
• Bipolar Disorder	35	(1.0)	32	(1.3)	69	(1.1)
• Bulimia	50	(1.4)	13	(0.5)	66	(1.1)
• Depression	422	(11.7)	137	(5.6)	568	(9.3)
• Insomnia	139	(3.9)	52	(2.1)	196	(3.2)
• Other sleep disorder	74	(2.0)	44	(1.8)	122	(2.0)
• Obsessive Compulsive Disorder	67	(1.9)	32	(1.3)	101	(1.6)
• Panic attacks	189	(5.2)	51	(2.1)	243	(4.0)
• Phobia	31	(0.9)	19	(0.8)	53	(0.9)
• Schizophrenia	0	(0.0)	12	(0.5)	14	(0.2)
• Substance abuse/addiction	22	(0.6)	32	(1.3)	57	(0.9)
• Other addiction	4	(0.1)	23	(0.9)	29	(0.5)
• Other mental health condition	41	(1.1)	31	(1.3)	75	(1.2)
Have you ever been diagnosed (by a professional) with depression?						
• Yes	656	(18.2)	243	(10.0)	908	(14.9)
Within the last 12 months, have any of the following been difficult to handle for you?						
• Academics	2,342	(64.3)	1,335	(54.2)	3,708	(60.3)
• Career-related issue	1,391	(38.2)	785	(31.9)	2,193	(35.7)
• Death of a family member or friend	599	(16.5)	258	(10.5)	865	(14.1)
• Family problems	1,198	(33.0)	504	(20.5)	1,719	(28.0)
• Other relationships	1,705	(47.1)	888	(36.2)	2,617	(42.7)
• Finances	1,498	(41.4)	793	(32.3)	2,312	(37.7)
• Health problem of a family member or friend	838	(23.1)	361	(14.7)	1,210	(19.7)
• Personal appearance	1,407	(39.0)	501	(20.4)	1,930	(31.5)
• Personal health issue	1,031	(28.5)	419	(17.0)	1,468	(23.9)
• Other	513	(14.4)	279	(11.5)	798	(13.2)
Have you ever received psychological or mental health services from any of the following?						
• Counselor/Psychologist	1,293	(35.8)	572	(23.3)	1,864	(30.8)
• Psychiatrist	467	(13.0)	263	(10.7)	738	(12.1)
• Other medical provider	575	(16.0)	187	(7.6)	768	(12.6)
• Clergy	229	(6.4)	144	(5.9)	379	(6.2)
Have you ever received psychological or mental health services from your current college/university's Counseling or Health Service?						
• Yes	567	(15.7)	220	(9.0)	794	(13.0)
If in the future you were having a personal problem that was really bothering you, would you consider seeking help from a mental health professional?						
• Yes	2,614	(72.2)	1,501	(61.2)	4,154	(67.8)

*All "yes" responses included: *Yes, diagnosed but not treated; Yes, treated with medication; Yes, treated with psychotherapy; Yes, treated with mediation and therapy; Yes, other treatment.*

disorder, schizophrenia, substance abuse/addiction, other addictions, and other mental health conditions with more frequency than women in this sample.

Altogether, 17.1 percent of students in this sample reported being diagnosed and/or treated for one or more mental health conditions in the last 12 months. A total of 7.7 percent of students experienced one condition, 4.6 percent experienced two conditions, 2.3 percent experienced three conditions, and 2.5 percent experienced four or more comorbidities. Nearly three-fourths (73.4 percent) of those students reporting any comorbidity (i.e., diagnosis and/or treatment with more than one mental health condition) were diagnosed or treated with anxiety in the last 12 months. Further, of the 542 students in this sample who reported experiencing 2 or more co-occurring conditions, 275 (50.7 percent) reported both anxiety and depression in the last 12 months.

In the overall sample, students reported that a variety of life events and experiences were difficult for them to handle in the last 12 months. As Figure 4.1 displays, only 17.5 percent of students reported no difficulty with any of the ten indicators specified. Correspondingly, 83.5 percent of students reported at least one difficulty in the last 12 months, 53.8 percent reported two or more difficulties in the last 12 months, and 39.7 percent reported three or more difficulties in the last 12 months. The five reported the most often were: academics (60.3 percent); other (i.e., non-family) relationships (42.7 percent); finances (37.7 percent); career-related issue (35.7 percent); and personal appearance (31.5 percent). Death of a family member or friend (14.1 percent) and "other" (13.2 percent) were reported with the lowest frequencies in this sample. It should be noted that college women endorsed each of these items with greater frequencies than college men in this sample. The biggest differences in proportions

across gender were for personal appearance (39.0 percent vs. 20.4 percent); family

problems (33.0 percent vs. 20.5 percent), and personal health issues (28.5 percent vs. 17.0

percent). When summed, women reported an average of 3.45 difficult life

events/experiences in the last 12 months and men reported an average of 2.47 difficult

life events in the last 12 months. The difference between the means was statistically

significant (t = 15.74, df = 5,899, p<.001).
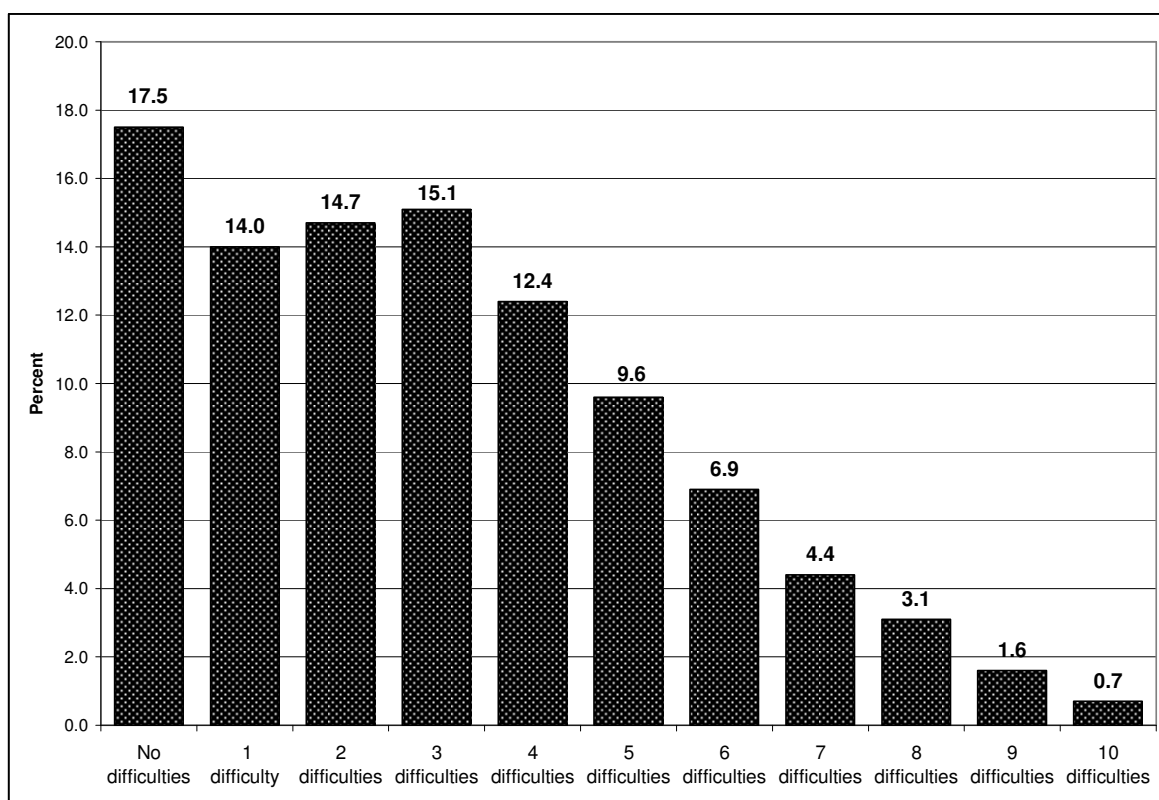
Figure 4.1. Number of difficulties experienced in the last 12 months in the revised
ACHA-NCHA student sample (valid n = 5,958).

As evidenced in Table 4.2, many students in this study reported receiving mental

health services from a variety of providers in their lifetime. In total, 36.4 percent of

students (41.7 percent of females and 29.2 percent of males) had ever received care from

at least of one the following types of mental health providers: a counselor or

psychologist, psychiatrist, other medical provider, and/or a clergy member. Nearly one in three (30.8 percent) had ever received services from a counselor or psychologist, 12.1 percent reported receiving services from a psychiatrist, 12.6 percent reported receiving services from other medical providers, and 6.2 percent reported receiving mental health services from a clergy member. Of the 2,209 students who reported seeking help from any of the aforementioned providers, 52.8 percent reported ever receiving help from only one type of mental health provider provided, 29.6 percent reported receiving help from two types of mental health providers, and 17.6 percent reported receiving help from three or more provider types in their lifetimes.

In addition to ever receiving mental health services from a variety of mental health providers, 13.0 percent of this sample reported receiving mental health services from their current college/university health or counseling services. Again, female students endorsed each of these items with greater frequency than male students. It should be highlighted that when asked, "If in the future you were having a personal problem that was really bothering you, would you consider seeking help from a mental health provider?" more than 2 in 3 students reported "yes" (67.7 percent of overall sample; 72.2 percent of females; and 61.2 percent of males).

Table 4.3 provides information about the proportion of students who reported ever injuring themselves, seriously considering suicide, and attempting suicide in the last 12 months. Altogether, 6.0 percent (n = 368) of students reported seriously considering suicide in the last 12 months and less than one percent (n = 49) reported an actual attempt. Students who identify as bi- or multiracial reported suicide ideation with greater frequency (9.7 percent) than other races in this sample, though this estimate is derived

from only a very small sample (n = 36). In addition to suicide ideation and attempts, 5.6

percent (n = 310) of this sample reported deliberate self-harm in the last 12 months.

These rates appear to remain relatively constant across year in school, and females (5.8

percent) report this behavior with greater frequency than males (4.0 percent), though

differences are not examined for statistical significance in this sample.

Table 4.3. Proportion of students who reported intentionally injuring selves, seriously considering suicide, and attempting suicide in the last 12 months on the revised ACHA-NCHA by various demographic characteristics.

	Intentionally injured self		Seriously considered suicide		Attempted suicide	
	n	(valid %)	n	(valid %)	n	(valid %)
Gender						
• Female (*n*=3,642)	210	(5.8)	233	(6.4)	30	(0.8)
• Male (*n*=2,469)	97	(4.0)	132	(5.4)	18	(0.7)
Race						
• White (*n*=4,149)	209	(5.1)	237	(5.7)	21	(0.5)
• Black (*n*=270)	14	(5.2)	17	(6.3)	4	(1.5)
• Hispanic or Latino/a (*n*=264)	8	(3.1)	10	(3.8)	1	(0.4)
• Asian/Pacific Islander (*n*=908)	41	(4.5)	51	(5.6)	13	(1.4)
• Bi/Multiracial (*n*=375)	26	(7.0)	36	(9.7)	6	(1.6)
• Other (*n*=124)	7	(5.8)	11	(9.0)	2	(1.7)
Year in School						
• 1st year undergraduate (*n*=1,216)	67	(5.5)	68	(5.6)	10	(0.8)
• 2nd year undergraduate (*n*=1,009)	49	(4.9)	56	(5.6)	4	(0.4)
• 3rd year undergraduate (*n*=1,360)	72	(5.3)	95	(7.0)	14	(1.0)
• 4th year undergraduate (*n*=1,043)	45	(4.3)	60	(5.7)	10	(1.0)
• 5th year or more undergraduate (*n*=326)	17	(5.0)	24	(7.4)	1	(0.3)
• Graduate or professional (*n*=1,085)	50	(4.7)	54	(5.0)	7	(0.7)
TOTAL (N=6,216)	310	(5.6)	368	(6.0)	49	(0.8)

Research question 1A

Research question 1A asked about the prevalence of a variety of mental health

conditions in this sample by demographic characteristics such as gender, race/ethnicity,

and year in school. Table 4.4 below presents past year prevalence estimates (diagnosis

and/or treatment in the last 12 months) for anxiety as well as past year and ever

prevalence estimates for depression for each of these groups.

Table 4.4. Percent of students who reported diagnosis/treatment with anxiety in the last 12 months, diagnosis/treatment with depression in the last 12 months, and ever being diagnosed with depression on the revised ACHA-NCHA by various demographic characteristics.

	Diagnosed or treated with anxiety in last 12 months		Diagnosed or treated with depression in last 12 months		Ever diagnosed with depression	
	n	(valid %)	n	(valid %)	n	(valid %)
Gender						
• Female (*n*=3,642)	387	(10.7)	422	(11.7)	656	(18.2)
• Male (*n*=2,469)	123	(5.0)	137	(5.6)	243	(10.0)
Race						
• White (*n*=4,149)	408	(9.9)	423	(10.3)	693	(16.9)
• Black (*n*=270)	12	(4.5)	18	(6.7)	28	(10.5)
• Hispanic or Latino/a (*n*=264)	15	(5.7)	20	(7.6)	30	(11.5)
• Asian/Pacific Islander (*n*=908)	29	(3.2)	44	(4.9)	59	(6.6)
• Bi/Multiracial (*n*=375)	31	(8.3)	38	(10.2)	61	(16.5)
• Other (*n*=124)	14	(11.4)	19	(15.4)	25	(20.5)
Year in School						
• 1st year undergraduate (*n*=1,216)	56	(4.6)	62	(5.1)	115	(9.6)
• 2nd year undergraduate (*n*=1,009)	65	(6.5)	66	(6.6)	102	(10.2)
• 3rd year undergraduate (*n*=1,360)	124	(9.2)	149	(11.0)	238	(17.6)
• 4th year undergraduate (*n*=1,043)	101	(9.7)	114	(11.0)	162	(15.8)
• 5th year or more undergraduate (*n*=326)	42	(14.6)	50	(15.4)	80	(24.9)
• Graduate or professional (*n*=1,085)	114	(10.6)	113	(10.5)	195	(18.2)
TOTAL (N=6,216)	521	(8.5)	568	(9.3)	908	(14.9)

Estimates of diagnosis/treatment with depression and anxiety in the last 12

months as well as ever being diagnosed with depression appear to generally increase

throughout a student's undergraduate education, particularly from the 2nd to the 3rd year.

Fifth year or more undergraduates in this sample endorse each of these items with greater

frequencies than most other groups and graduate students report each of these conditions with similar frequencies to 3rd and 4th year undergraduates. In general, White, bi/multiracial, and students of other racial identities endorse each of these items with higher frequencies than Black, Hispanic or Latino/a, and Asian/Pacific Islander students. For all groups except graduate students, the 12-month prevalence of depression is higher than for anxiety; in graduate students, the prevalence estimates are nearly identical for the two conditions (10.5 percent vs. 10.6 percent, respectively). As would certainly be expected, the "ever" diagnosed with depression prevalence estimates are larger for each group than the 12-month prevalence estimates. The groups reporting the highest percentages of ever being diagnosed with depression include fifth year or more undergraduates (24.9 percent), those identifying with an "other" racial or ethnic identity (20.5 percent), females (18.2 percent), graduate students (18.2 percent), and third year undergraduates (17.6 percent).

In addition to evaluating the prevalence estimates for anxiety and depression by a variety of demographic characteristics, it was also of interest to this researcher to determine if any differences existed by institution. Because the sample for the revised ACHA-NCHA is complex in that it drew students from seven PSIs, some variability in differences in proportions could possibly be attributed to institution. Hence, for each of the estimates above (i.e., 12-month prevalence of anxiety, 12-month prevalence of depression, and ever diagnosed with depression), cross-tabs (*school × experienced mental health condition*) were produced and chi-square and Phi values were estimated. Table 4.5 presents the results of these analyses. No statistically significant differences existed across school for diagnosis with depression ever or within in the last 12 months. Though

statistically significant differences existed across school for diagnosis or treatment with

anxiety in the last 12 months, the Phi coefficient is quite small (.054), indicating the

effect size of this difference is of little practical significance (see Chapter Three for more

information in regards to Phi coefficients, effect sizes, and practical significance). Hence,

for these mental health conditions, school of attendance appeared to have little effect on

prevalence estimates.

Table 4.5. Examining differences in the percent of students reporting diagnosis/treatment with anxiety in the last 12 months, diagnosis/treatment with depression in the last 12 months, and ever being diagnosed with depression on the revised ACHA-NCHA by institution.

	Percent of overall sample*	Range of percents across PSIs	Chi-square**	p-value	Cramers V
Diagnosed or treated with anxiety in last 12 months	8.5	5.5 - 10.3	17.782	<.05	.054
Diagnosed or treated with depression in last 12 months	9.3	7.8 - 10.2	5.079	>.05	.029
Ever diagnosed with depression	14.9	12.7 - 17.7	12.298	>.05	.045

*Overall N=6,216
**Each chi-square analysis had 6 degrees of freedom.

Though it is somewhat alarming that nearly a fifth to a quarter of some student

demographic groups have been diagnosed and/or treated with depression in their

lifetimes, it is promising that many students have also sought help from mental health

providers. As mentioned above, females reported receiving care from any mental health

provider (i.e., counselor or psychologist, psychiatrist, other medical provider, and clergy)

at higher frequencies that males. Table 4.6 provides an overview of the percentages of

students who report ever seeking help from a number of mental health providers, ever

seeking help from their current college or university health center, and a willingness to consider seeking help in the future by a variety of demographic characteristics. When all provider types are collapsed, the following percentages of students report receiving care from at least one type of mental health provider by race and by year in school:

By race:
- White: 40.2%
- Black: 32.0%
- Hispanic or Latino/a: 40.5%
- Asian or Pacific Islander: 17.7%
- Bi- or Multiracial: 41.7%
- Other: 40.2%

By year in school:
- 1st year undergraduate: 29.4%
- 2nd year undergraduate: 31.0%
- 3rd year undergraduate: 38.0%
- 4th year undergraduate: 38.7%
- 5th year or more undergrad: 50.2%
- Graduate or professional: 41.5%

Table 4.6. Percent of student respondents reporting ever seeking mental health services from a variety of providers and their current college/university, and a willingness to consider seeking help in the future by various demographic characteristics.

	By provider type				From current college's counseling or health center	Con-sider seeking help in the future
	Couns-elor	Psych-iatrist	Other medical provider	Clergy		
Gender						
• Female (*n*=3,642)	35.8	13.0	16.0	6.4	15.7	72.2
• Male (*n*=2,469)	23.3	10.7	7.6	5.9	9.0	61.2
Race						
• White (*n*=4,149)	34.5	14.0	14.1	6.4	13.2	71.2
• Black (*n*=270)	24.5	7.4	7.5	9.3	10.8	55.7
• Hispanic or Latino/a (*n*=264)	32.8	6.9	11.9	8.4	14.9	65.9
• Asian/Pacific Islander (*n*=908)	12.3	4.2	6.4	3.7	10.0	58.4
• Bi/Multiracial (*n*=375)	35.1	12.9	12.8	6.8	16.2	65.9
• Other (*n*=124)	33.3	20.5	20.0	7.4	17.5	62.0
Year in School						
• 1st year undergraduate (*n*=1,216)	24.3	7.9	9.4	4.8	4.8	59.4
• 2nd year undergraduate (*n*=1,009)	26.8	8.9	9.5	4.7	11.2	63.5
• 3rd year undergraduate (*n*=1,360)	30.6	12.9	14.1	7.5	14.7	67.1
• 4th year undergraduate (*n*=1,043)	31.9	13.0	13.6	6.5	17.4	71.4
• 5th year or more undergrad (*n*=326)	44.4	21.2	19.2	8.4	24.2	69.5
• Graduate/professional (*n*=1,085)	36.1	15.0	14.0	6.6	14.4	78.1
TOTAL (N=6,216)	30.3	11.9	12.4	6.1	13.0	67.8

Fifth year or more undergraduates report ever receiving mental health care from a

counselor (44.4 percent), psychiatrist (21.2 percent), and their current college's

counseling or health center (24.2 percent) at rates that appear higher than all combined

students (30.3 percent, 11.9 percent, and 13.0 percent, respectively). Contrarily, Asian

students report receiving care from counselors (12.3 percent), psychiatrists (6.4 percent),

and their university's counseling and health centers (10.0 percent) at lower percentages

than average. Black students and Hispanic or Latino/a students report ever seeking

mental health services from clergy and religious officials (9.3 percent and 8.4 percent,

respectively) at rates higher than average (6.1 percent).

As described previously in Table 4.2, approximately two-thirds of students in this

sample report a willingness to seek help in the future for a problem that was really

bothering them. These estimates vary tremendously by various demographic

characteristics, as provided in Table 4.5. For example, graduate students report a

willingness to seek help at the highest rates (78.1 percent) in this sample. In general,

willingness to seek help increases steadily from 1st year undergraduates (59.4 percent) to

4th year undergraduates (71.4 percent) in this sample. Among those of various racial and

ethnic identities, White students report a willingness to seek help with the highest

frequency (71.2 percent), and Asian/Pacific Islander (58.4 percent) and Black (55.7

percent) students with the lowest.

Figure 4.2 on the following page displays levels of stress reported by students in

this sample according to a variety of demographic characteristics. In total, 1.5 percent of

students in this sample reported no stress, 7.9 percent reported less than average stress,

40.5 percent reported average stress, 41.0 percent reported above average stress, and 9.1

	Female	Male		White	Black	Hisp./ Latino	Asian/ PI	Bi/Multi-racial	Other		1st year	2nd year	3rd year	4th year	5th year or more	Grad-uate		TOTAL
Tremendous	10.2	7.5		9.1	9.7	7.2	7.4	11.8	14.6		6.5	7.2	9.6	9.2	10.2	12.9		9.1
> Average	44.6	35.5		42.2	33.7	41.1	37.1	44.0	34.1		36.5	37.0	40.1	41.0	50.6	47.7		41.0
Average	40.0	41.5		40.7	41.6	39.5	41.3	37.0	42.3		43.9	45.1	41.8	40.7	35.2	32.9		40.5
< Average	4.7	12.7		7.0	12.7	11.0	11.2	5.1	5.7		11.2	9.0	6.9	8.0	3.7	5.5		7.9
None	0.5	2.8		1.0	2.2	1.1	3.0	2.1	3.3		2.0	1.6	1.6	1.2	0.3	1.1		1.5

Figure 4.2. Levels of stress reported by students on the revised ACHA-NCHA by various demographic characteristics.

percent reported tremendous stress. In general, females report higher levels of experienced stress than males. More than half of females (54.8 percent) report "above average" or "tremendous" stress while males report these stress levels with lower frequencies (44.0 percent). Racial and ethnic differences also appear to exist in students' self-reported experiences of stress in the last 12 months. Asian/Pacific Islander and Hispanic/Latino students report lower levels of stress than other racial and ethnic groups, while bi/multiracial and students of other racial and ethnic identities report the highest levels of perceived stress. It also appears that self-reported stress levels increase by year in school, with freshman reporting the lowest levels of stress and fifth-year or more undergraduate and graduate students reporting the highest.

Again, because of the nature of the complex sampling, it was of interest to this researcher to determine if, in addition to differences in stress level observed by various demographic characteristics, any differences between institutions appeared to exist in students' self-reported levels of stress. When analyzed initially in a 7×5 contingency table (*institution × level of stress*), results were statistically significant ($\chi^2 = 84.826$, df =24, p < .001). However, Cramer's V (a measure of association similar to the Phi coefficient that is used in tables larger than 2×2) was quite small at .059, indicating once again that these observed significant differences may be limited in their practical significance. When further examined to determine if any factors such as school size, location, or institutional type appeared to impact self-reported stress levels, no patterns appeared to emerge. Subsequently, in this sample, institution did not appear to dramatically impact students' self-reported stress levels and issues of complex sampling will be ignored in further analyses.

In summary, the preceding section of this Chapter provided an overview of the frequencies with which students endorsed each of the 44 mental health indicators on the revised ACHA-NCHA. Information for each indicator is available in total and by gender, and information on many indicators was further provided by race/ethnicity and year in school. Though many patterns appeared to emerge, differences in these endorsements were not assessed for their statistical significance by various demographic characteristics because no hypotheses were being evaluated for this research question. Further, many cell sizes were quite small once they were stratified by these various demographics. However, statistical significance of how endorsements varied by institution was of interest for some indicators, particularly those that were to be entered into future regression models (e.g., diagnosis or treatment with depression in the last 12 months, stress level). Because no practically significant differences existed among the four indicators evaluated by institution, institutional characteristics were not considered in subsequent analyses, and as is the case in numerous studies utilizing complex sampling (Garson, n.d.), all future analyses considered only individual characteristics (e.g., race/ethnicity) as opposed to institutional characteristics (e.g., PSI size, PSI location).

The following section of results provides an overview of how student endorsements of mental health indicators on the revised ACHA-NCHA differ from comparable indicators on the original ACHA-NCHA.

Research Question 2

The second research question of interest in this study asked, "What differences exist in student response patterns between comparable original ACHA-NCHA mental

health items and revised ACHA-NCHA mental health items?" Because the literature

demonstrated that even small changes in survey questions have produced significant

changes in survey results, it was hypothesized that changes in question formatting and

response options between the original and revised ACHA-NCHA mental health

indicators would result in significant differences in patterns of student responses on

comparable items across versions of the survey.

　　　　To evaluate this hypothesis, 2×2 contingency tables were created for each of the

comparable survey items in which one variable in the table was the *version of the survey*

(revised or original) and the other was the *pattern of student responses*. Data from each

version of the survey were collapsed to form two response categories for comparison

(e.g., "In the last 12 months, Not in the last 12 months" on the revised survey vs. "In the

last school year, Not in the last school year" on the original survey). Data were analyzed

according to the plan set forth in Table 3.4, and Pearson's chi-square (χ^2) for independent

proportions and Phi coefficients were calculated for each of the 13 comparisons of

interest. Because the sample sizes across both versions of the survey were very large—

more than 6,000 respondents in each—there was sufficient power to detect even very

small effects. Hence, only those comparisons that resulted in Phi coefficients of 0.1 or

higher (a standard for small effects; Miles & Gilbert, 2005) were considered practically

significant and in support of hypothesis 2.

　　　　Table 4.7 provides information in regards to the five direct comparisons made

across versions for indicators on the negative affect scale and for the two direct

comparisons made for suicide ideation and attempts. For the purposes of these analyses,

responses from the revised ACHA-NCHA, the stem of which read "Have you ever:" were

Table 4.7. Comparisons of response proportions for comparable negative affect scale indicators across versions of the ACHA-NCHA.

	Original Reponses *Within the last school year:*	Revised Responses *Within the last 12 months:*	Chi-square value, df, p-value, Phi
Have you felt things were hopeless	*Never* (n = 2,412; 39.8%) • Never *In the last school year:* (n = 3,641; 60.2%) • 1-2 times, 3-4 times, 5-6 times, 7-8 times, 9-10 times, 11 or more times (combined)	*Never* (n = 3,287; 53.6%) • Never, Not in the last 12 months (combined) *In the last 12 months:* (n = 2,840; 46.4%) • In the last 12 months, In the last 30 days, In the last 2 weeks (combined)	χ^2: 232.900 df: 1 p-value: <.001 Phi: -.138
Have you felt overwhelmed by all you had to do	*Never* (n = 427; 7.1%) • Never *In the last school year:* (n = 5,625; 92.9%) • 1-2 times, 3-4 times, 5-6 times, 7-8 times, 9-10 times, 11 or more times (combined)	*Never* (n = 795; 12.9%) • Never, Not in the last 12 months (combined) *In the last 12 months:* (n= 5,350; 87.1%) • In the last 12 months, In the last 30 days, In the last 2 weeks (combined)	χ^2: 117.010 df: 1 p-value: <.001 Phi: -.098
Have you felt exhausted (not from physical activity)	*Never* (n = 577; 9.5%) • Never *In the last school year:* (n = 5,472; 90.5%) • 1-2 times, 3-4 times, 5-6 times, 7-8 times, 9-10 times, 11 or more times (combined)	*Never* (n = 1,096; 17.8%) • Never, Not in the last 12 months (combined) *In the last 12 months:* (n = 5,045; 82.2%) • In the last 12 months, In the last 30 days, In the last 2 weeks (combined)	χ^2: 177.657 df: 1 p-value: <.001 Phi: -.121
Have you felt very sad	*Never* (n = 1,363; 22.6%) • Never *In the last school year:* (n = 4,678; 77.4%) • 1-2 times, 3-4 times, 5-6 times, 7-8 times, 9-10 times, 11 or more times (combined)	*Never* (n = 2,134; 34.8%) • Never, Not in the last 12 months (combined) *In the last 12 months:* (n = 4,001; 65.2%) • In the last 12 months, In the last 30 days, In the last 2 weeks (combined)	χ^2: 222.082 df: 1 p-value: <.001 Phi: -.135
Have you felt so depressed that it was difficult to function	*Never* (n = 3,451; 57.1%) • Never *In the last school year:* (n = 2,594; 42.9%) • 1-2 times, 3-4 times, 5-6 times, 7-8 times, 9-10 times, 11 or more times (combined)	*Never* (n = 4,192; 68.3%) • Never, Not in the last 12 months (combined) *In the last 12 months:* (n = 1,946; 31.7%) • In the last 12 months, In the last 30 days, In the last 2 weeks (combined)	χ^2: 163.631 df: 1 p-value: <.001 Phi: -.116
Have you seriously considered suicide	*Never* (5,545; 91.6%) • Never *In the last school year:* (n = 508; 8.4%) • 1-2 times, 3-4 times, 5-6 times, 7-8 times, 9-10 times, 11 or more times (combined)	*Never* (n = 5,775; 94.0%) • Never, Not in the last 12 months (combined) *In the last 12 months:* (n = 368; 6.0%) • In the last 12 months, In the last 30 days, In the last 2 weeks (combined)	χ^2: 26.385 df: 1 p-value: <.001 Phi: -.047
Have you attempted suicide	*Never* (n = 5,966; 98.9%) • Never *In the last school year:* (n = 69; 1.1%) • 1-2 times, 3-4 times, 5-6 times, 7-8 times, 9-10 times, 11 or more times (combined)	*Never* (n = 6,086; 99.2%) • Never, Not in the last 12 months (combined) *In the last 12 months:* (n = 49; 0.8%) • In the last 12 months, In the last 30 days, In the last 2 weeks (combined)	χ^2: 3.763 df: 1 p-value: .052 Phi: -.018

collapsed to *Never in the last 12 months* (response options: Never, Not in the last 12 months) and *In the last 12 months* (response options: In the last 12 months, In the last 30 days, In the last 2 weeks) for each variable. Responses from the original ACHA-NCHA, the stem of which read "Within the last school year, how many times have you:" were collapsed to *Never in the last school year* (response option: Never) and *In the last school year* (response options: 1-2 times, 3-4 times, 5-6 times, 7-8 times, 9-10 times, 11 or more times) for each variable. Hence, the two variables on the 2×2 contingency tables were Version of the survey (original and revised) and Experienced in the last year (never and ever). It should be noted that the time periods on these comparison are not directly comparable—the original version of the survey asks students to report on their experiences within the last school year while the revised version of the survey asks students to report on their experiences within the last 12 months. The impact this difference in time period has on interpretation of results is discussed in Chapter Five.

There are many noteworthy findings presented in Table 4.7. Each of the comparisons made between the five indicators on the negative affect scale was statistically significant when alpha was equal to .05. Pearson's chi-square values for these differences ranged from 117.010 for feeling overwhelmed in the last year to 232.900 for feeling things were hopeless in the last year. All phi coefficients were negative (Φ range: -.098 to -.138), which in these analyses indicate that "ever experienced" in the last year was lower for each of these indicators on the revised survey when compared to the original. For example, on the original version of the survey, 60.2 percent of respondents indicated they had felt hopeless one or more times in the last school year; on the revised survey, which utilized different response options, 46.4 percent of students reported ever

feeling hopeless in the last 12 months. Phi values are highest for "felt things were hopeless" (Φ = -.138) and "felt very sad" (Φ = -.135), thereby indicating the strength of the relationship between the version of the survey and the pattern of student responses was highest (i.e., the largest differences existed) for these two indicators. Not surprisingly, because Phi is a function of chi-square and sample size, chi-square estimates were highest for these as well. The smallest difference existed for students' reports of feeling overwhelmed by all they had to do in the last year. On the original version of the survey, 92.9 percent of students reported feeling overwhelmed in the last school year one or more times, and on the revised version of the survey 87.1 percent of students reported feeling overwhelmed in the last 12 months (χ^2: 117.010, df =1, p<.001, Φ = -.098).

Much smaller differences between versions of the survey existed in student response patterns for the variables measuring consideration of suicide and suicide attempts. On the original version of the survey, 8.4 percent of students reported seriously considering attempting suicide at least once in the last school year. This percentage was significantly higher than the 6.0 percent of students who reported ever seriously considering suicide in the last 12 months on the revised survey (χ^2: 26.385, df =1, p<.001, Φ = -.047), yet the Phi coefficient indicates the effect size of this difference is very small. Furthermore, no statistically significant differences existed in the proportion of students who reported attempting suicide one or more times in the last school year on the original survey (1.1 percent) and ever attempting suicide in the last 12 months on the revised survey (0.8 percent).

Table 4.8 presents comparisons for prevalence estimates for a variety of mental

health conditions across versions of the survey. As described in Chapter Three, the

questions of interest from the original survey to be used for these comparisons were:

- **Q43:** Have you ever been diagnosed with depression? (*Yes, No*)

- **Q47:** Within the last school year, have you had any of the following? (*No, Yes*)
    - B. Anorexia
    - C. Anxiety disorder
    - D. Bulimia
    - F. Depression

- **Q48:** Within the last school year, have you had any of the following? (*No, Yes*)
    - M. Substance abuse problem

The questions of interest to be used for the revised survey for these analyses were:

- **NQ31:** Have you ever been diagnosed (by a professional) with depression? (*No, Yes*)

- **NQ30:** Within the last 12 months, have you been diagnosed or treated by a professional for any of the following? (*No; Yes, diagnosed but not treated; Yes, treated with medication; Yes, treated with psychotherapy; Yes, treated with medication and psychotherapy; Yes, other treatment*)
    - P. Anorexia
    - Q. Anxiety
    - E. Bulimia
    - F. Depression
    - M. Substance abuse or addiction (alcohol or other drugs)

In order to create comparisons between items across versions of the survey, all "yes"

responses were collapsed for the revised version. It should be noted that the time periods

on these comparison are not directly comparable—the original version of the survey asks

students to report on their experiences with these conditions within the last school year

while the revised version of the survey asks students to report on their experiences within

the last 12 months. Again, the impact this difference in time period has on interpretation

of results is discussed in Chapter Five. Other differences between items (e.g., diagnosis

Table 4.8. Comparisons of response proportions for comparable mental illness prevalence items across versions of the ACHA-NCHA.

	Original Reponses *Within the last school year…:*	Revised Responses *Within the last 12 months…:*	Chi-square value, df, p-value, Phi
Depression (ever diagnosed)	Have you ever been diagnosed with depression? • *No* (n = 5,254; 87.0%) • *Yes* (n = 785; 13.0%)	Have you ever been diagnosed (by a professional) with depression? • *No* (n = 5,190; 85.1%) • *Yes* (n = 908; 14.9%)	χ^2: 9.042 df: 1 p-value: .003 Phi: .027
Anorexia	Have you had any of the following: anorexia? • *No* (n = 5,959; 98.6%) • *Yes* (n = 85; 1.4%)	Have you been diagnosed or treated by a professional with any of the following: anorexia? • *No* (n = 6,091; 99.1%) • *Yes* (n = 57, 0.9%) Yes, diagnosed but not treated; Yes, treated with medication; Yes, treated with psychotherapy; Yes, treated with medication and psychotherapy; Yes, other treatment (combined)	χ^2: 6.080 df: 1 p-value: .014 Phi: -.022
Anxiety	Have you had any of the following: anxiety disorder? • *No* (n = 5,327; 88.3%) • *Yes* (n = 708; 11.7%)	Have you been diagnosed or treated by a professional with any of the following: anxiety? • *No* (n = 5,609; 91.5%) • *Yes* (n = 521; 8.5%) Yes, diagnosed but not treated; Yes, treated with medication; Yes, treated with psychotherapy; Yes, treated with medication and psychotherapy; Yes, other treatment (combined)	χ^2: 34.985 df: 1 p-value: <.001 Phi: -.054
Bulimia	Have you had any of the following: bulimia? • *No* (n = 5,890; 97.8%) • *Yes* (n = 131; 2.2%)	Have you been diagnosed or treated by a professional with any of the following: bulimia? • *No* (n = 6,607; 98.9%) • *Yes* (n = 66; 1.1%) Yes, diagnosed but not treated; Yes, treated with medication; Yes, treated with psychotherapy; Yes, treated with medication and psychotherapy; Yes, other treatment (combined)	χ^2: 29.172 df: 1 p-value: <.001 Phi: -.048
Depression	Have you had any of the following: depression? • *No* (n = 5,092; 84.3%) • *Yes* (n = 951; 15.7%)	Have you been diagnosed or treated by a professional with any of the following: depression? • *No* (n = 5,566; 90.7%) • *Yes* (n = 568; 9.3%) Yes, diagnosed but not treated; Yes, treated with medication; Yes, treated with psychotherapy; Yes, treated with medication and psychotherapy; Yes, other treatment (combined)	χ^2: 116.976 df: 1 p-value: <.001 Phi: -.098
Substance abuse	Have you had any of the following: substance abuse problem? • *No* (n = 5,863; 97.1%) • *Yes* (n = 173; 2.9%)	Have you been diagnosed or treated by a professional with any of the following: Substance abuse or addiction (alcohol or other drugs)? • *No* (n = 6,066; 99.1%) • *Yes* (n = 57; 0.9%) Yes, diagnosed but not treated; Yes, treated with medication; Yes, treated with psychotherapy; Yes, treated with medication and psychotherapy; Yes, other treatment (combined)	χ^2: 61.340 df: 1 p-value: <.001 Phi: -.071

as opposed to diagnosis by a professional for the ever diagnosed with depression question) are also discussed.

All prevalence items were statistically significant when alpha is equal to .05. Comparisons of past year prevalence indicators across versions produced chi-square values that ranged from 6.080 for anorexia (Φ = -.022) to 116.976 for depression (Φ = -.098). Again, because the phi values are negative, results indicate that prevalence in the last year was lower for each of these indicators on the revised survey when compared to the original. Furthermore, each of the phi values are relatively low (i.e., < .100), which suggests that though the differences between versions of the survey are statistically significant, they may be limited in the degree to which they are of practical significance.

When student responses from "Have you ever been diagnosed with depression?" (*Yes*, *No*) on the original version of the survey were compared to "Have you ever been diagnosed (by a professional) with depression?" (*No*, *Yes*) on the revised version, significantly more students responded "yes" on the revised version (χ^2 = 9.042, df = 1, p = .003). The phi value associated with this comparison (.027) is very low, indicating the observed differences across versions of the survey for this mental health indicator were very small and not of great practical significance.

Hypothesis 2

In summary, results of these chi-square tests for independence indicate that 12 of the 13 mental health indicators evaluated produced statistically significant differences in student response patterns across versions of the ACHA-NCHA. Of the 13 comparisons made, however, only 4 produced Phi coefficients of 0.1 or higher, each of which was for an indicator on the negative affect scale. Past-year prevalence estimates for a variety of

mental health conditions and for ever diagnosed with depression, though statistically significant, produced very small effects. Correspondingly, because of Phi cutoff values set forth in Chapter Three, Hypothesis 2, which stated "Changes in question formatting and response options between the original and revised ACHA-NCHA mental health indicators will result in significant differences in patterns of student responses on comparable items across versions of the survey" is only partially supported. Issues related to statistical significance and practical significance of these reported differences will be discussed further in Chapter Five.

Research Question 3

The third research question of interest in this study was designed to assess the level to which the mental health indicators on the revised ACHA-NCHA were reliable. To evaluate the two hypotheses associated with this question, psychometric properties associated with the negative affect scale were evaluated using a principal components analysis (PCA) to determine scale dimensionality followed by a reliability analysis to determine internal consistency.

The negative affect scale on the revised ACHA-NCHA is comprised of eight indicators (felt things were hopeless, felt overwhelmed by all you had to do, felt exhausted—not from physical activity, felt very lonely, felt very sad, felt so depressed it was difficult to function, felt overwhelming anxiety, felt overwhelming anger), each of which asked students to report whether they had ever experienced the feeling through the response options of "Never, Not in the last 12 months, In the last 12 months, In the last 30 days, In the last 2 weeks."

Figure 4.3 presents the scree plot that was produced when results from the negative affect scale were analyzed using a principal components analysis (PCA) with direct oblimin rotation. As is common practice with PCA, results from the scree plot and the Kaiser-Guttman rule (i.e., eigenvalues greater than one) were analyzed in tandem to determine the underlying scale dimensionality. Because the point of inflexion on the scree plot suggests there is a sharp decline in the amount of variance explained by the second and other components, and because the eigenvalue for the first principal component was 4.587 while that of the second component was only 1.003, results suggested that the scale was likely unidimensional in nature. The first principal component explained a total of 57.3 percent of the variance in the eight underlying variables. Inter-item correlations for the eight items are presented in Table 4.9. The indicators "felt very lonely" and "felt very sad" correlated the most highly among items ($r = .789$) in this sample, and "felt overwhelmed by all you had to do" and "felt overwhelming anger" correlated with the lowest coefficient ($r = .328$) in this sample.
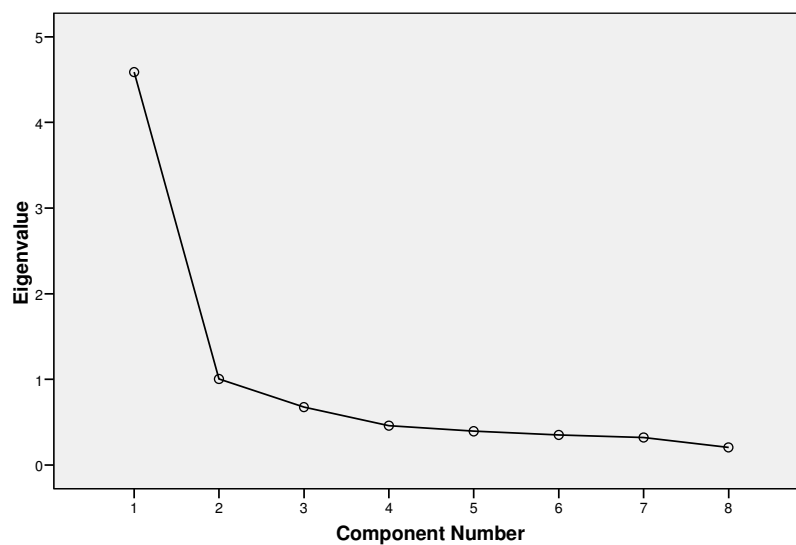
Figure 4.3. The scree plot generated from a principal components analysis using direct oblimin rotation for the eight indicators on the revised ACHA-NCHA negative affect scale.

Table 4.9. Correlation matrix for items on the negative affect scale from the revised ACHA-NCHA. Question stem reads "In the last 12 months, have you felt…" and response options were "*Never; Not in the last 12 months; In the last 12 months; In the last 30 days; In the last 2 weeks*."

	Things were hopeless	Overwhelmed	Exhausted	Very lonely	Very sad	So depressed it was difficult to function	Overwhelming anxiety	Overwhelming anger
Things were hopeless	1.000	.408	.412	.592	.614	.619	.556	.507
Overwhelmed	.408	1.000	.651	.418	.445	.327	.436	.328
Exhausted	.412	.651	1.000	.463	.480	.368	.443	.353
Very lonely	.592	.418	.463	1.000	.789	.569	.502	.471
Very sad	.614	.445	.480	.789	1.000	.620	.548	.523
So depressed it was difficult to function	.619	.327	.368	.620	.620	1.000	.629	.565
Overwhelming anxiety	.556	.436	.443	.548	.548	.629	1.000	.602
Overwhelming anger	.507	.328	.353	.523	.523	.565	.602	1.000

Table 4.10 presents factor loadings and communalities for each of the eight

indicators on the single extracted principal component. Factor loadings are essentially

correlations of each indicator with the final created principal components (Pett et al.,

2003). Generally speaking, an indicator variable is said to load sufficiently on a principal

component if its factor loading is .400 or higher (Dunteman, 1989). Communalities (h^2

estimates) are estimates of the variance in each indicator explained by the principal

components of interest (i.e., the common variance), and in the case of a single extracted

principal component are equal to squared factor loadings (Pett et al.). As Table 4.10

suggests, factor loadings for each of the indicators on the negative affect scale were

relatively high for each of these eight indicators (range: .642 - .841), indicating strong

correlations between each of these standardized items and the principal component that

was created from these variables. The highest loadings were for "felt very sad" (.841) and

"felt very lonely" (.805), and the lowest were for "felt overwhelmed by all you had to do"

(.642) and "felt exhausted (not from physical activity)" (.672). Correspondingly, the

variance explained by the principal components for each of these variables is highest for

"felt very sad" (h^2 = .708) and is lowest for "felt overwhelmed by all you had to do" (h^2 =

.413).

Table 4.10. Factor loadings and communalities for each item on the negative affect scale.

	Factor loading	Communality (h^2)
Felt things were hopeless	.786	.618
Felt overwhelmed	.642	.413
Felt exhausted	.672	.451
Felt very lonely	.805	.648
Felt very sad	.841	.708
Felt so depressed it was difficult to function	.787	.619
Felt overwhelming anxiety	.783	.613
Felt overwhelming anger	.720	.518

Because all indicators loaded sufficiently onto the principal component extracted, and because all results indicated the negative affect scale on the revised ACHA-NCHA was unidimensional, the eight indicators were next analyzed using a scale analysis for their internal consistency. Table 4.11 presents results of this analysis. Item means on the scale ranged from 2.09 ("felts so depressed it was difficult to function") to 3.88 ("felt overwhelmed from all you had to do"). Higher means indicate greater average recency for each of the indicators. Standard deviations of each of the indicators were relatively constant (range: 1.266 – 1.429). Corrected item-total correlations (i.e., correlations with each indicator with the total scale when the item is removed from the scale) ranged from .447 for "felt overwhelming anger" to .691 for "felt very sad" in this sample.

Table 4.11. Various psychometric characteristics of each item on the revised ACHA-NCHA negative affect scale.

	Mean	Standard deviation	Corrected item-total correlation	Squared multiple correlation	Cronbach's alpha if item deleted
Felt things were hopeless	2.51	1.397	.700	.515	.876
Felt overwhelmed	3.88	1.266	.557	.466	.889
Felt exhausted	3.75	1.360	.585	.484	.887
Felt very lonely	3.00	1.407	.722	.649	.874
Felt very sad	3.06	1.378	.769	.691	.869
Felt so depressed it was difficult to function	2.09	1.277	.701	.562	.877
Felt overwhelming anxiety	2.56	1.429	.700	.538	.876
Felt overwhelming anger	2.31	1.324	.626	.447	.883

Hypotheses 3A and 3B

The internal consistency for these eight indicators was estimated to be .893, which was greater than the standard cutoff value of acceptable internal consistency (.700, see Nunnally & Bernstein, 1994) set forth in Chapter Three. Hence, Hypothesis 3A, which stated, "Measurement scales on the revised ACHA-NCHA will demonstrate sufficient levels of internal consistency in the pre-test sample," was supported.

As is often the case in a scale analysis, these indicators were further examined to determine if the estimated alpha would increase if the indicator was removed from the scale. Results of this analysis, which are presented in the last column of Table 4.11, suggest Cronbach's alpha would decrease if any of the indicators were removed (range of corrected α estimates: .869 – .889), thereby suggesting each of the indicators contributes to the internal consistency of the scale. Therefore, Hypothesis 3B, which stated, "Reliability will not increase if any indicators are removed from measurement scales on the revised ACHA-NCHA in the pre-test sample," was supported.

Research Question 4

The fourth research question in this study asked, "To what extent are the mental health indicators on the revised ACHA-NCHA valid?" Evidence of validity of data collected from the revised survey was determined primarily through criterion-related validity and construct-related validity assessment, common forms of validity assessment in survey research. Results associated with Hypothesis 4A, which examined criterion-related evidence of validity, and Hypothesis 4B, which examined construct-related evidence of validity are presented below.

Hypothesis 4A

Criterion-related evidence of validity is established when results from a survey are compared to that of a "gold-standard" of measurement. For example, when a students' self-report of diagnosis with depression is confirmed by examination of a medical record that indicates its truth, criterion-related evidence of validity is provided. Although medical records are perhaps the most stringent comparison for the purposes of

evaluating validity of 12-month prevalence estimates of mental health conditions, those data are unavailable for the students surveyed in this sample. Consequently, another commonly used type of criterion-related validity assessment was used—that which compares prevalence estimates collected from the revised survey to those collected from a nationally generalizable survey instrument.

Table 4.12 provides 12-month prevalence estimates for a variety of mental health conditions by gender, age group, and in total collected from the revised ACHA-NCHA and the National Co-morbidity Study Replication ([NCS-R], Kessler et al., 2005). Though not all subjects surveyed in the NCS-R attended colleges or universities, and as such are not directly comparable to the study population being evaluated in this dissertation, these estimates are currently seen as the "gold standard" of mental disorder prevalence. Chapter Three provides a more comprehensive overview of the NCS-R and its methodology.

As is seen in Table 4.12, ACHA-NCHA past-year prevalence estimates vary greatly in the degree to which they differ from NCS-R estimates. For many of the mental health disorders, prevalence estimates were relatively similar: for example, the 12-month prevalence estimate for depression in the ACHA-NCHA sample was very similar to the 12-month prevalence estimate of "any mood disorder" in the NCS-R sample (ACHA-NCHA: 9.3 percent, NCS-R: 9.7 percent), as were the 12 month prevalence estimate for Obsessive Compulsive Disorder in both samples (ACHA-NCHA: 1.6 percent, NCS-R: 1.2 percent). A summary of findings is provided below with key relevant information:

Table 4.12. Twelve-month prevalence estimates for a variety of mental disorders by gender and age: A comparison of results from the revised ACHA-NCHA and the National Co-morbidity Study Replication (NCS-R). All values are percents.

Mental Disorder for which prevalence is estimated	Survey	Gender		Age		Total (N = 6,216)*
		Females (N = 3,642)*	Males (N = 2,469)*	18-29 years (N = 5,755)*	30-44 years (N = 324)*	
Anorexia	ACHA-NCHA	1.0	0.7	0.9	0.6	0.9
	NCS-R[1]	0.0	0.0	0.0	0.0	0.0
Anxiety	ACHA-NCHA	10.7	5.0	8.3	11.5	8.5
	NCS-R[2]	3.4	1.9	2.0	3.5	2.7
Attention Deficit and Hyperactivity Disorder (ADHD)	ACHA-NCHA	2.2	3.7	2.9	2.2	2.9
	NCS-R[3]	3.9	4.3	3.9	4.2	4.1
Bipolar disorder	ACHA-NCHA	1.0	1.3	1.0	1.9	1.1
	NCS-R[4]	2.8	2.9	4.7	3.5	2.8
Bulimia	ACHA-NCHA	1.4	0.5	1.1	0.9	1.1
	NCS-R[5]	0.5	0.1	-	-	0.3
Depression	ACHA-NCHA	11.7	5.6	8.9	14.6	9.3
	NCS-R[6]	11.6	7.7	12.9	11.9	9.7
Obsessive Compulsive Disorder (OCD)	ACHA-NCHA	1.9	1.3	1.6	2.2	1.6
	NCS-R[7]	1.8	0.5	1.5	1.4	1.2
Panic attacks	ACHA-NCHA	5.2	2.1	3.8	5.6	4.0
	NCS-R[8]	3.8	1.6	2.8	3.7	2.7
Phobia	ACHA-NCHA	0.9	0.8	0.8	1.3	0.9
	NCS-R[9]	12.2	5.8	10.3	9.7	9.1
Substance abuse or addiction (Alcohol or other drugs)	ACHA-NCHA	0.6	1.3	0.9	1.2	0.9
	NCS-R[10]	0.7	2.2	3.9	1.2	1.4

*   All N values are for the revised ACHA-NCHA sample.
1.   Estimates presented from Hudson et al. (2007). Findings suggest no 12-month prevalence of anorexia. Lifetime prevalence was estimated at 0.6% (females: 0.9%, males 0.3%)
2.   Estimates presented from Kessler et al. (2005). NCS-R estimates are for 12-month prevalence of "Generalized anxiety disorder". 12-month prevalence of "Any anxiety disorder" was 19.1% (females:23.4%, males: 14.3%)
3.   Estimates presented from Kessler et al. (2005).
4.   Estimates presented from Kessler et al. (2005). NCS-R estimates are for 12-month prevalence of "Bipolar I-II-sub disorders"
5.   Estimates presented from Hudson et al. (2007).
6.   Estimates presented from Kessler et al. (2005). NCS-R estimates are for 12-month prevalence of "Any mood disorder" which includes major depressive disorder, dysthymia, and bipolar I-II-sub disorders. 12-month prevalence of major depressive disorder alone was 6.8% (females: 8.6%, males: 4.9%)
7.   Estimates presented from Kessler et al. (2005).
8.   Estimates presented from Kessler et al. (2005). NCS-R estimates are for "Panic disorder".
9.   Estimates presented from Kessler et al. (2005). NCS-R estimates are for "Specific phobia".
10. Estimates presented from Kessler et al. (2005). NCS-R estimates are for "Drug abuse with/without dependence". 12-month prevalence of alcohol abuse with/without dependence was 3.1% (females: 1.8%, males: 4.5%)

- *Anorexia:* According to Hudson and colleagues (2007), the NCS-R found the 12-month prevalence of anorexia nervosa to be 0.0 percent in their sample. However, they did find the lifetime prevalence of anorexia to be 0.6 percent (0.9 percent among women in their sample and 0.3 percent among men). The ACHA-NCHA found the 12-month prevalence among college students to be 0.9 percent, with higher prevalence estimates among women than men and younger students compared to older students; these observed patterns in the revised ACHA-NCHA sample are similar to gender patterns found in lifetime prevalence estimates from the NCS-R.

- *Anxiety:* According to Kessler et al. (2005), the NCS-R found the 12-month prevalence of generalized anxiety disorder to be 2.7 percent in their study, with higher prevalence estimates for women (3.4 percent) than men (1.9 percent) and older adults (3.5 percent) than younger (2.0 percent). It should be noted that the NCS-R appears to conceptualize anxiety differently than the ACHA-NCHA in that Generalized Anxiety Disorder appears to be a specific diagnosis in a series of many anxiety-related diagnoses, including Obsessive Compulsive Disorder, phobias, and Post-Traumatic Stress Disorder. When all anxiety-related disorders are included in the NCS-R estimates, the 12-month prevalence of anxiety is 19.1 percent (females: 23.4 percent, males: 14.3 percent; 18-29 years: 22.3 percent, 30-44 years: 22.7 percent). The ACHA-NCHA found a 12-month prevalence estimate of anxiety of 8.5 percent, with higher percentages for women than men and older adults than younger. Though estimates vary tremendously across the two surveys, the patterns across various demographic groups are similar.

- *Attention Deficit and Hyperactivity Disorder (ADHD):* The NCS-R (Kessler et al., 2005) found a 12-month prevalence estimate of ADHD of 4.1 percent, with a higher prevalence for males than females and a slightly higher prevalence for older adults than younger. The ACHA-NCHA found a 12-month prevalence estimate of ADHD of 2.9 percent with a higher prevalence reported among males

than females. However, unlike the NCS-R, the ACHA-NCHA sample reported a higher prevalence among younger adults (2.9 percent) than older adults (2.2 percent).

- ***Bipolar disorder:*** The estimate provided for 12-month prevalence of bipolar disorder was 2.8 percent according to the NCS-R (Kessler et al., 2005), with a higher prevalence among females than males and among younger adults than older adults. The 12-month prevalence estimates for bipolar disorder from the revised ACHA-NCHA is lower (1.1 percent of the entire sample) than that reported on the NCS-R. Moreover, patterns reported by the respondents of the ACHA-NCHA differed from those reported by the NCS-R respondents: the ACHA-NCHA sample reported higher prevalence among males (1.3 percent) than females (1.0 percent) and among older (1.9 percent) rather than younger (1.0 percent) adults.

- ***Bulimia:*** The 12-month prevalence estimate of bulimia reported by the NCS-R was 0.3 percent (Hudson et al., 2007), with a higher prevalence estimate among females than males. Estimates were unavailable by age group, though the authors did report an estimated age of onset of 19.7 years. The 12-month prevalence estimate of bulimia in the ACHA-NCHA sample was 1.1 percent, with higher prevalence estimates for females than males and for younger adults compared to older adults.

- ***Depression:*** The NCS-R reported a 12-month prevalence of "any mood disorder" at 9.7 percent (Kessler et al., 2005). It should be noted that this included major depressive disorder, dysthymia, and bipolar I and II disorders. The NCS-R reported a higher 12-month prevalence of mood disorders among women than men and among younger adults than older adults. The ACHA-NCHA reported a very similar prevalence estimate of depression (9.3 percent), though this percentage does not include bipolar disorder as it did in the NCS-R. Like the NCS-R, the ACHA-NCHA reported higher prevalence estimates for depression

for women than for men, but unlike the NCS-R, estimates were higher in the revised sample for older as opposed to younger adults.

- *Obsessive Compulsive Disorder (OCD):* Kessler and colleagues (2005) reported a 12-month prevalence of OCD at 1.2 percent, with higher rates among women. The revised ACHA-NCHA found a 12-month prevalence estimate of 1.6 percent, and like the NCS-R, a higher estimate for women (1.9 percent) than for men (1.3 percent). Both studies found comparable prevalence estimates among younger and older adults.

- *Panic attacks:* The NCS-R (Kessler et al., 2005) found a 12-month prevalence of panic disorder of 2.7 percent, with rates more than double for women (3.8 percent) than for men (1.6 percent). Prevalence of panic disorder was higher for older adults than younger adults in the NCS-R sample. The ACHA-NCHA found a 12-month prevalence estimate of panic attacks—it should be noted that that ACHA-NCHA asked respondents about panic attacks as opposed to panic disorder—to be 4.0 percent. Patterns similar to those found on the NCS-R for panic disorder emerged on the ACHA-NCHA for panic attacks: women reported higher 12-month prevalence estimates than men (5.2 vs. 2.1 percent, respectively), and older students reported higher prevalence than younger students (5.6 vs. 3.8 percent, respectively).

- *Phobia:* Nearly one in ten respondents on the NCS-R (9.1 percent; Kessler et al., 2005) reported experiencing a specific phobia within the past 12 months. Women (12.2 percent) reported higher prevalence estimates of phobias than men (5.8 percent), and estimates did not vary greatly by age group. The revised ACHA-NCHA sample found a much lower 12-month prevalence estimate of phobias at 0.9 percent. Estimates did not vary greatly by gender or by age group in this sample. These dramatic differences in estimates are explored in more depth in Chapter Five.

- *Substance abuse or addiction:* The NCS-R (Kessler et al., 2005) reported a 12-month prevalence of drug abuse with/without dependence of 1.4 percent, with higher estimates for men than women and for younger as opposed to older adults. It should be noted that the NCS-R asked respondents separately about alcohol abuse with or without dependence. The 12-prevalence estimates for alcohol abuse with or without dependence were 3.1 percent for the overall sample, 1.8 percent for females, 4.5 percent for males, 7.1 percent for 18-29 year-olds, and 3.3 percent for 30-44 year-olds. The ACHA-NCHA asked about substance abuse or addiction (alcohol or other drugs). Overall, 0.9 percent of the revised survey sample reported being diagnosed or treated for substance abuse or addiction in the last 12 months. Like the NCS-R, 12-month prevalence estimates were higher for men than for women, but unlike the NCS-R, they were lower for younger students (0.9 percent) than for those aged 30-44 years (1.2 percent).

In summary, estimates of 12-month prevalence were most comparable across the ACHA-NCHA and the NCS-R for ADHD, depression on the ACHA-NCHA compared to "any mood disorder" on the NCS-R, OCD, panic attacks on the ACHA-NCHA compared to "panic disorder" on the NCS-R, and substance abuse or addiction (alcohol or other drugs) on the ACHA-NCHA compared to "drug abuse with/without dependence" on the NCS-R. Estimates were less comparable for anorexia, anxiety on the ACHA-NCHA compared to "generalized anxiety disorder" on the NCS-R, bipolar disorder, and bulimia. The largest observed differences in 12-month prevalence estimates were for phobia on the ACHA-NCHA compared to "specific phobia" on the NCS-R.

Patterns in responses (e.g., lower prevalence for females than males, higher prevalence for younger adults than older adults) also varied to the extent they were similar. Comparable patterns across gender were observed for anxiety, ADHD, bipolar disorder, bulimia, depression, OCD, panic attacks, and substance abuse but not for

anorexia or phobias. Very few patterns were similar between age groups: only for anxiety and panic attacks did similar trends between younger adults and older adults (i.e., younger adults reported lower prevalence estimates than older adults) appear on both the ACHA-NCHA and the NCS-R. For all other estimates (i.e., anorexia, ADHD, bipolar disorder, depression, OCD, phobias, and substance abuse), the patterns between 18-29 year olds and 30-44 year-olds differed between ACHA-NCHA and NCS-R results. It should be noted, however, that only a very small proportion (5.3 percent) of the revised ACHA-NCHA sample was aged 30- 44 years. Therefore, these results must be interpreted with caution as discussed in more depth in Chapter Five.

Because of the varying degrees to which results and patterns observed in data collected from the revised ACHA-NCHA differed from the NCS-R, Hypothesis 4A, which stated, "As a result of comparison with data collected from other surveys of mental health among college-aged students, the mental health indicators on the revised ACHA-NCHA will demonstrate evidence of criterion-related validity in the pre-test sample," is only partially supported.

Hypothesis 4B

In addition to assessing criterion-related evidence of validity, this study also utilized three regression (two linear and one logistic) models to assess evidence of construct-related evidence of validity to evaluate Hypothesis 4B. The exact specifications for each of these models are provided in Chapter Three. Construct-related evidence of validity is assessed by simultaneously examining convergent and discriminant validity. Linear regression models, through the examination of R^2, b, and beta values, provided an opportunity to examine evidence of these construct-related types of validity for

continuous outcome variables. A logistic regression model, through the examination of b and Exp (b) values, provided a similar opportunity for a dichotomous outcome.

Results from all models evaluated in these analyses are presented in Tables 4.13, 4.14, and 4.15. Each of the three models controlled for demographic characteristics in the first block of the model. Because age, year in school, gender, and race/ethnicity have all been well-documented as characteristics on which mental health status may vary, each was entered into the regression models. Males served as the reference group for gender because females have been well-documented as the at-risk group for a variety of mental health concerns. Additionally, two dummy variables were created for race with White students as the reference group. The first included Black, Hispanic/Latino, and Asian/Pacific Islander students and the second included Bi/multiracial and "other" racial/ethnic identities. These dummy variables were chosen because, as presented above in the section addressing Research Question 1A, similar patterns appeared to emerge on many variables for Black, Hispanic/Latino, and Asian/Pacific Islander students while others emerged among those students identifying as Bi/multiracial and or "other" racial/ethnic identities.

Table 4.13 presents the results for the first linear regression model (i.e., Model A) specified in Chapter 3, that which regressed recency of negative affect on self-reported stress level when controlling for various demographic characteristics. The first block of the model (Model A1), that which regressed recency of negative affect (i.e., a factor score created from the eight indicators of the negative affect scale) on only demographic characteristics found that the five demographic predictors of interest explained 5.0 percent of the variance in the outcome. Though this model was better than no predictors

Table 4.13. Ordinary least squares regression predicting recency of negative affect from stress when controlling for various demographic characteristics in the revised ACHA-NCHA sample (Model A, N=5,785).

Variable	B	(95% CI of B)	Standard Error of B	β	t	p-value
Model A1: *Demographics Only*[1]						
• Constant	.055	(-.068, .179)	.063	-	.875	.381
• Age	-.017	(-.023, -.010)	.003	-.082	-5.064	<.001
• Year in school	.024	(.005, .042)	.009	.041	2.526	.012
• Gender : *Female (vs. Male)*	.398	(.347, .449)	.026	.196	15.310	<.001
• Race: *Black/Hispanic/Asian (vs. White)*	-.093	(-.153, -.033)	.031	-.040	-3.039	.002
• Race: *Bi/Multiracial/Other (vs. White)*	.241	(.148, .334)	.047	.066	5.091	<.001
Model A2: *Demographics and Stress Level*[2]						
• Constant	-1.730	(-1.866, -1.593)	.070	-	-24.838	<.001
• Age	-.020	(-.026, -.015)	.003	-.101	-7.153	<.001
• Year in school	-.007	(-.024, .009)	.008	-.013	-.886	.376
• Gender: *Female (vs. Male)*	.237	(.192, .282)	.023	.117	10.258	<.001
• Race: *Black/Hispanic/Asian (vs. White)*	-.012	(-.065, .040)	.027	-.005	-.460	.645
• Race: *Bi/Multiracial/Other (vs. White)*	.189	(.108, .271)	.041	.052	4.565	<.001
• Level of stress	.589	(.561, .616)	.014	.484	41.990	<.001

1. $R^2_{\text{Model A1}}$ = .050 (adjusted R^2 = .049), F change from model with no predictors = 60.312 (df = 5, 5779; p < .001)
2. $R^2_{\text{Model A2}}$ = .272 (adjusted R^2 = .271), F change from Model A1 = 1763.193 (df = 1, 5778; p < .001)

Table 4.14. Ordinary least squares regression predicting level of stress from top 5 reported difficulties in the last 12 months when controlling for various demographic characteristics in the revised ACHA-NCHA sample (Model B, N=5,858).

Variable	B	(95% CI of B)	Standard Error of B	β	t	p-value
Model B1: *Demographics Only*[1]						
• Constant	3.034	(2.933, 3.134)	.051	-	59.245	<.001
• Age	.006	(.001, .012)	.003	.038	2.383	.017
• Year in school	.055	(.040, .070)	.008	.114	7.129	<.001
• Gender : *Female (vs. Male)*	.277	(.235, .319)	.021	.166	13.025	<.001
• Race: *Black/Hispanic/Asian (vs. White)*	-.145	(-.194, -.096)	.025	-.075	-5.824	<.001
• Race: *Bi/Multiracial/Other (vs. White)*	.077	(.001, .154)	.039	.026	1.993	.046
Model B2: *Demographics and Difficulties in the last 12 months*[2]						
• Constant	1.641	(1.514, 1.768)	.065	-	25.341	<.001
• Age	.010	(.005, .015)	.002	.061	4.142	<.001
• Year in school	.047	(.033, .061)	.007	.098	6.516	<.001
• Gender: *Female (vs. Male)*	.172	(.133, .211)	.020	.103	8.583	<.001
• Race: *Black/Hispanic (vs. White)*	-.152	(-.197, -.107)	.023	-.079	-6.582	<.001
• Race: *Bi/Multiracial/Other (vs. White)*	.011	(-.060, .081)	.036	.003	.293	.769
• Difficult to handle: Academics	.396	(.353, .440)	.022	.236	17.953	<.001
• Difficult to handle: Career-related issue	.154	(.110, .199)	.023	.090	6.804	<.001
• Difficult to handle: Other relationships	.165	(.124, .206)	.021	.099	7.854	<.001
• Difficult to handle: Finances	.109	(.067, .152)	.022	.065	5.009	<.001
• Difficult to handle: Personal appearance	.131	(.085, .176)	.023	.074	5.655	<.001

1. $R^2_{Model\ B1}$ = .053 (adjusted R^2 = .052), F change from model with no predictors = 65.760 (df = 5, 5852; p < .001)
2. $R^2_{Model\ B2}$ = .197 (adjusted R^2 = .196), F change from Model B1 = 210.208 (df = 5, 5847; p < .001)

Table 4.15. Binary logistic regression predicting diagnosis and/or treatment with depression in the last 12 months from stress and recency of negative affect when controlling for various demographic characteristics in the revised ACHA-NCHA sample (Model C, N=5,763).

Variable	B	Standard Error of B	Exp(B)	(95% CI of Exp(B))	Wald	p-value
Model C1: *Demographics Only[1]*						
• Constant	-3.697	.200	.025	-	340.965	<.001
• Age	.027	.009	1.027	(1.009, 1.045)	8.715	.003
• Year in school	.109	.032	1.115	(1.048, 1.187)	11.729	.001
• Gender : *Female (vs. Male)*	.823	.105	2.276	(1.854, 2.795)	61.636	<.001
• Race: *Black/Hispanic/Asian (vs. White)*	-.635	.129	.530	(0.412, 0.682)	24.280	<.001
• Race: *Bi/Multiracial/Other (vs. White)*	.138	.158	1.148	(0.842, 1.565)	.758	.384
Model C2: *Demographics, Stress Level, and Recency of Negative Affect[2]*						
• Constant	-5.045	.322	.006	-	245.485	<.001
• Age	.044	.010	1.045	(1.025, 1.065)	19.710	<.001
• Year in school	.094	.034	1.098	(1.028, 1.173)	7.777	.005
• Gender: *Female (vs. Male)*	.563	.109	1.756	(1.418, 2.175)	26.558	<.001
• Race: *Black/Hispanic (vs. White)*	-.645	.134	.524	(0.403, 0.682)	23.263	<.001
• Race: *Bi/Multiracial/Other (vs. White)*	-.142	.167	.868	(0.625, 1.204)	.721	.396
• Level of stress	.239	.073	1.270	(1.100, 1.465)	10.670	.001
• Recency of negative affect	.871	.060	2.388	(2.124, 2.685)	212.322	<.001

1. Omnibus test for model coefficient (from no predictors to Model 3A): χ^2 = 139.659 (df = 5, p < .001); Cox & Snell R^2 = .024, Nagelkerke R^2 = .052
2. Omnibus test for model coefficients (from Model 3A to Model 3B): χ^2 = 374.653 (df = 2, p < .001); Cox & Snell R^2 = .085, Nagelkerke R^2 = .185

($F_{5,5779} = 60.312$, p < .001), it was not nearly as useful in predicting the outcome of interest as the model that also included self-reported level of stress (Model A2). When stress was included in the model, the predictors of interest explained a total of 27.2 percent of the variance in recency of negative affect; subsequently, this model was significantly more useful in predicting the outcome of interest (R^2 change = .222, F change = 1763.2, p < .001).

Though all demographic variables were statistically significant predictors of negative affect in Model A1, when stress level was added to Model A2, year in school and one of the race dummy variables (Black/Hispanic/Asian vs. White) were no longer significant predictors. In Model A2, that which included all predictors of interest, Beta values indicate that level of stress was the predictor of the most relative importance (β = .484). These results also found that Biracial, Multiracial, and students identifying as "Other" racial identities had higher levels of negative affect when controlling for all other variables in the model (b = .189, p < .001), as did females when compared to males (b = .237, p < .001). When controlling for all other predictors in the model as age increased by one year, the recency of negative affect measure decreased by -.020 standard deviations (because recency of negative affect is a factor score, it is a standardized value and its units are in standard deviations). Ultimately, level of stress was a statistically significant predictor of recency of negative affect when controlling for the demographic variables of interest, thereby confirming relationships expected according to literature and theory and demonstrating construct-related evidence of validity of these data.

Table 4.14 presents the results for the second linear regression model (i.e., Model B) specified in Chapter 3, which regressed self-reported stress level on the top five

difficulties experienced by students in the revised ACHA-NCHA sample when

controlling for various demographic characteristics. The first block of the model, that

which regressed self-reported stress level on only demographic characteristics found that

the six demographic predictors of interest explained 5.3 percent of the variance in the

outcome. Though this model was better than no predictors ($F_{5,5852} = 65.760$, $p < .001$), it

was not nearly as useful in predicting the outcome of interest as the model that also

included the five most commonly reported difficulties (i.e., academics, career-related

issue, other relationships, finances, and personal appearance) (Model B2). When

experiences with these difficulties (each was dichotomous with "was not difficult to

handle in the last 12 months" as the reference group) were included in the model, the

predictors of interest explained a total of 19.7 percent (adjusted $R^2 = .196$) of the variance

in self-reported stress levels; subsequently, this model was significantly more useful in

predicting the outcome of interest (R^2 change = .144, F change = 210.208, $p < .001$) than

the model with solely demographic predictors.

Though all demographic variables were statistically significant predictors of

negative affect in Model B1, when experiences with various difficulties within the last 12

months were added to Model B2, one of the race dummy variables (Bi/Multiracial/Other

vs. White) was no longer a significant predictor. In Model B2, that which included all

predictors of interest, Beta values indicate that difficulty with academics in the last 12

months was the most important predictor in the model ($\beta = .236$), followed by difficulty

with other (i.e., non-familial) relationships in the last 12 months ($\beta = .099$), and year in

school ($\beta = .098$). When controlling for all demographic characteristics and each other,

each of the five most commonly reported difficulties was a statistically significant

predictor of self-reported stress levels in this sample. This indicates that each of the five

most reported difficulties in the last 12 months had its own bearing on self-reported

stress.

These results also found that Black, Hispanic or Latino/a, and Asian/Pacific

Islander students had lower levels of stress than White students when controlling for all

other variables in the model (b = -.152, p < .001), and females had higher levels of stress

than males when controlling for all other variables (b = .237, p < .001). Hence, results

indicate the relationship between experiencing difficult life events and stress level may

differ by race and by gender. Ultimately, experiencing each of the five most common

difficulties in the last 12 months predicted students' self-reported level of stress when

controlling for the demographic variables of interest, thereby confirming relationships

expected according to literature and theory and demonstrating construct-related evidence

of validity of these data.

The third model tested to evaluate Hypothesis 4B utilized logistic regression to

predict diagnosis and or/treatment with depression in the last 12 months from level of

stress and recency of negative affect when controlling for demographic variables of

interest. Results from this model (Models 3A and 3B as specified in Chapter Three) are

presented in Table 4.15. The demographic variables alone contributed significantly to the

model as the -2 log likelihood (-2LL) value decreased significantly (χ^2 change = 139.659,

df = 5, p < .001) from the model with no predictors (-2LL = 3557.494) to the model with

the five demographic predictors (-2LL = 3417.835). In this model (Model 3A), all

predictors except one of the race dummy variables (Biracial/Multiracial/Other racial

identity vs. White) were significantly related to the log odds of the outcome variable,

diagnosis and/or treatment with depression in the last 12 months. When holding all other demographic predictors in the model constant, odds of being diagnosed and/or treated with depression significantly increased by each year of age (OR = 1.027, 95% CI: 1.009, 1.045, p = .003) and each year in school (OR = 1.115, 95% CI: 1.048, 1.187, p = .001). Further, females were 2.276 times as likely as males to report 12-month prevalence of depression (95% CI of OR: 1.854, 2.795, p < .001) and students who identify as Black, Hispanic or Latino/a, or Asian/Pacific Islander were 0.530 times as likely as White students to report diagnosis and/or treatment with depression in the last 12 months (95% CI of OR: 0.412, .0682, p < .001), holding all else constant.

Model C2 incorporated each of the above demographic variables into the predictive model as well as self-reported level of stress in the last 12 months and recency of negative affect, a standardized factor score derived from the eight indicators that comprise the negative affect scale. Adding these two independent variables significantly increased the predictive ability of the model (-2LL = 3043.182, χ^2 = 374.653, df = 2, p < .001). Once again, the only statistically insignificant predictor in the model was the race dummy variable of Biracial/Mutiracial/Other students compared to White students. Comparable patterns among demographic predictors (i.e., increased odds of depression in the last 12 months by year of age, year in school, and for females compared to males; decreased odds of depression in the last 12 months for Black, Hispanic/Latino, and Asian/Pacific Islander students compared to Whites) remained when controlling for stress level and recency of negative affect. When holding all demographics and recency of negative affect constant, each one unit increase in self-reported stress level made students 1.270 times as likely (95% CI of OR: 1.100, 1.465, p = .001) to report diagnosis and/or

treatment with depression in the last 12 months. Similarly, when holding all demographics and level of stress constant, each one standard deviation increase in recency of negative affect made students 2.388 times as likely (95% CI of OR: 2.124, 2.685, $p < .001$) to report diagnosis and/or treatment with depression in the last 12 months. Hence, level of stress and recency of negative affect both served as statistically significant predictors of depression in the last 12 months when controlling for one another and for race, year in school, age, and gender, thereby confirming relationships expected according to literature and theory and demonstrating construct-related evidence of validity of these data.

Each of the three models used to evaluate evidence of construct-related validity in these data suggested expected relationships exist among constructs. As such, Hypothesis 4B, which stated, "As a result of analyses demonstrating expected theoretical relationships between variables, mental health indicators on the revised ACHA-NCHA will demonstrate evidence of construct-related validity in the pre-test sample," was supported by these findings.

Research Question 5

The fifth, and final, research question in this study asked, "How do mental health indicators on the revised ACHA-NCHA differ from those on the original ACHA-NCHA in terms of their reliability and validity?" Because a great deal of effort went into modifying, clarifying, and revising the original ACHA-NCHA based on feedback from customers and respondents, reviews of literature, and survey response theory, it was hypothesized that data collected from the mental health indicators on the revised

instrument would demonstrate more evidence of reliability and validity than data collected from the original instrument. To test this hypothesis, data collected from the negative affect scale were evaluated and compared for their internal consistency reliability across versions of the survey, and past year prevalence estimates were compared to one another and to estimates from the NCS-R to compare their criterion-related evidence of validity.

The negative affect scale on the original ACHA-NCHA contained the following five items, each of which began with the stem, "Within the last school year, how many times have you…" and had the response options of "Never, 1-2 times, 3-4 times, 5-6 times, 7-8 times, 9-10 times, 11 or more times":

- Felt things were hopeless
- Felt overwhelmed by all you had to do
- Felt exhausted (not from physical activity)
- Felt very sad
- Felt so depressed it was difficult to function

When data collected from the five items were entered into a PCA with oblique rotation, a one component solution appeared most appropriate. Correspondingly, the five items were examined using a reliability analysis. Cronbach's alpha for the five negative affect scale items on the original survey was .881, and alpha did not increase if any items were removed from the scale (see Table 4.16).

Table 4.16. Various psychometric characteristics of each item on the original ACHA-NCHA negative affect scale.

	Mean	Standard deviation	Corrected item-total correlation	Squared multiple correlation	Cronbach's alpha if item deleted
Felt things were hopeless	2.52	1.890	.727	.612	.852
Felt overwhelmed	4.32	2.033	.677	.590	.864
Felt exhausted	4.25	2.095	.681	.594	.864
Felt very sad	3.07	1.972	.804	.698	.833
Felt so depressed it was difficult to function	2.07	1.744	.695	.627	.861

Table 4.16 presents information about a variety of the psychometric properties of

the indicators of the negative affect scale from the original ACHA-NCHA. When

compared to results from Table 4.11, that which provides psychometric properties of the

negative affect scale from the original ACHA-NCHA negative affect scale, item means

for various indicators (e.g., felt things were hopeless, felt very sad), were very similar

across versions of the survey despite the changes to the item response options. In general,

standard deviations of the indicators were lower on the original version of the survey, and

corrected item-total correlations were higher. Ultimately, because the Cronbach's alpha

estimated from the indicators on the negative affect scale on the original version of the

survey (.881) was lower than that estimated from the indicators on the revised version of

the survey (.893), the negative affect scale constructed on the revised ACHA-NCHA

appeared to demonstrate higher levels of internal consistency reliability, thus supporting

Hypothesis 5.

The second aspect of Research Question 5 was interested in the differences of

criterion-related validity between the original and revised past-year prevalence estimates

of a variety of mental health conditions, including anorexia, bipolar disorder, bulimia,

depression, and substance abuse. As was the case in Research Question 4, data collected

from both versions of the survey were compared to the nationally generalizable estimates

collected from the NCS-R across gender and age groups. Table 4.17 presents these

results.

For all conditions on which there were comparable prevalence data available on

both versions of the ACHA-NCHA, estimates from the revised ACHA-NCHA appeared

to be closer to those provided on the NCS-R. For example, when diagnosis/treatment

Table 4.17. Twelve-month prevalence estimates for a variety of mental disorders by gender and age: A comparison of results from the revised ACHA-NCHA, the original ACHA-NCHA, and the National Co-morbidity Study Replication (NCS-R). All values are percents.

Mental Disorder for which prevalence is estimated	Survey	Gender		Age		Total* $(N_R = 6,216)$ $(N_O = 6,110)$
		Females* $(N_R = 3,642)$ $(N_O = 3,682)$	Males* $(N_R = 2,469)$ $(N_O = 2,352)$	18-29 years* $(N_R = 5,755)$ $(N_O = 5,655)$	30-44 years* $(N_R = 324)$ $(N_O = (303)$	
Anorexia	Revised ACHA-NCHA	1.0	0.7	0.9	0.6	0.9
	Original ACHA-NCHA	1.9	0.6	1.5	0.0	1.4
	NCS-R[1]	0.0	0.0	0.0	0.0	0.0
Anxiety disorder	Revised ACHA-NCHA	10.7	5.0	8.3	11.5	8.5
	Original ACHA-NCHA	14.2	8.0	11.5	16.8	11.7
	NCS-R[2]	3.4	1.9	2.0	3.5	2.7
Bulimia	Revised ACHA-NCHA	1.4	0.5	1.1	0.9	1.1
	Original ACHA-NCHA	3.1	0.7	2.3	0.3	2.2
	NCS-R[3]	0.5	0.1	-	-	0.3
Depression	Revised ACHA-NCHA	11.7	5.6	8.9	14.6	9.3
	Original ACHA-NCHA	17.7	12.6	15.1	26.7	15.7
	NCS-R[4]	11.6	7.7	12.9	11.9	9.7
Substance abuse or addiction (Alcohol or other drugs)	Revised ACHA-NCHA	0.6	1.3	0.9	1.2	0.9
	Original ACHA-NCHA	2.4	3.5	2.9	2.0	2.9
	NCS-R[5]	0.7	2.2	3.9	1.2	1.4

* N_R values are for the revised ACHA-NCHA total sample; N_O values are for the original ACHA-NCHA total sample.

1. Estimates presented from Hudson et al. (2007). Findings suggest no 12-month prevalence of anorexia. Lifetime prevalence was estimated at 0.6% (females: 0.9%, males 0.3%)

2. Estimates presented from Kessler et al. (2005). NCS-R estimates are for 12-month prevalence of "Generalized anxiety disorder". 12-month prevalence of "Any anxiety disorder" was 19.1% (females: 23.4%, males: 14.3%)

3. Estimates presented from Hudson et al. (2007).

4. Estimates presented from Kessler et al. (2005). NCS-R estimates are for 12-month prevalence of "Any mood disorder" which includes major depressive disorder, dysthymia, and bipolar I-II-sub disorders. 12-month prevalence of major depressive disorder alone was 6.8% (females: 8.6%, males: 4.9%)

5. Estimates presented from Kessler et al. (2005). NCS-R estimates are for "Drug abuse with/without dependence". 12-month prevalence of alcohol abuse with/without dependence was 3.1% (females: 1.8%, males: 4.5%)

with depression in the last 12 months on the revised ACHA-NCHA (9.3 percent) was compared to "had depression within the last school year" on the original ACHA-NCHA (15.7 percent) and then "any mood disorder" from the NCS-R (9.7 percent), estimates appear to be more congruent between the revised ACHA-NCHA and the NCS-R than the original ACHA-NCHA and the NCR-R. Generally speaking, past year prevalence estimates collected from the original ACHA-NCHA were higher than estimates collected from both the revised ACHA-NCHA and the NCS-R. As discussed previously in the results from Research Question 4, there are several differences in question wording between the ACHA-NCHA and the NCS-R, thereby limiting the extent to which direct comparisons can be made. This is further discussed in Chapter Five. However, it does appear that prevalence estimates collected from the revised ACHA-NCHA are more similar to those collected from the NCS-R, thereby generating further support for Hypothesis 5.

Ultimately, the negative affect scale on the revised ACHA-NCHA demonstrated evidence of higher internal consistency reliability than that on the original ACHA-NCHA, and past year prevalence estimates collected from the revised ACHA-NCHA were more similar to the NCS-R than those collected from the original ACHA-NCHA, thereby providing further evidence of criterion-related validity. Thus, Hypothesis 5, which states, "Mental health indicators on the revised ACHA-NCHA will demonstrate evidence of higher levels of reliability and validity than those on the original ACHA-NCHA," was supported by these findings.

CHAPTER V: DISCUSSION

Introduction

The first aim of this study was to examine the psychometric properties (i.e., reliability and validity) and patterns of student responses on the 44 mental health indicators on the revised ACHA-NCHA. The second was to evaluate the extent to which observed patterns in student responses differed on the 13 comparable mental health items across versions of the NCHA and to evaluate which version of the survey demonstrated higher evidence of reliability and validity. An experimental pre-test methodology, the utility of which is well-documented (e.g., Presser et al., 2004), was used to collect data from participating students and to evaluate the five specified research questions.

Ultimately, because of the investments made by the ACHA and this researcher in the development of the revised survey instrument, it was hoped and intended that data collected from the experimental pre-test of the revised survey would (1) independently demonstrate evidence of reliability and validity, (2) demonstrate higher evidence of reliability and validity than the original ACHA-NCHA when applicable, and (3) suggest this survey will be a useful and appropriate tool to measure college students' mental health status at PSIs across the country. Ultimately, results seem to support each of these intentions.

This chapter begins with a discussion of results for each research question and its hypotheses and follows with a summary of key findings (see Table 5.1). It then follows with an overview of the study's limitations and implications and concludes with directions for future research.

Table 5.1. Summary of key research findings.

Research Question/Hypothesis	Summary of Key Findings
Research Question 1. *What patterns of student responses emerge on each of the revised ACHA-NCHA mental health indicators?*	
• **Research Question 1A:** How do student endorsements of each item vary according to demographic characteristics such as gender, year in school, and race/ethnicity?	Difference in responses appear to exist on a variety of items, particularly by gender. Females appear to endorse most indicators with higher frequencies than males, though these results are not evaluated for statistical significance. Fifth-year undergraduates and bi/multiracial students also appear report a variety of conditions at higher levels.
Research Question 2. *What differences exist in student response patterns between comparable original ACHA-NCHA mental health items and revised NCHA mental health items?*	
• **Hypothesis 2:** Changes in question formatting and response options between the original and revised ACHA-NCHA mental health indicators will result in significant differences in patterns of student responses on comparable items across versions of the survey.	This hypothesis is partially supported. 12 of the 13 comparisons across versions are statistically significant, with lower frequencies on the revised survey for nearly all. Only 4 comparison resulted in a Phi coefficient of .100 or higher. Greater differences existed across versions for indicators of negative affect than for suicide ideation, past year prevalence of mental illnesses, and lifetime prevalence of depression. No differences emerged for suicide attempts.
Research Question 3. *To what extent are the mental health indicators on the revised ACHA-NCHA reliable?*	
• **Hypothesis 3A:** Measurement scales on the revised ACHA-NCHA will demonstrate sufficient levels of internal consistency in the pre-test sample.	This hypothesis is supported. The eight indicators of the negative affect scale have a reliability coefficient of .893, which exceeds the standard of .700.
• **Hypothesis 3B:** Reliability will not increase if any indicators are removed from measurement scales on the revised ACHA-NCHA in the pre-test sample.	This hypothesis is supported. When indicators are individually removed from the scale, Cronbach's alpha decreases to 869 - .889. Hence, alpha does not increase if any indicator is removed.
Research Question 4. *To what extent are the mental health indicators on the revised ACHA-NCHA valid?*	
• *Hypothesis 4A*: As a result of comparison with data collected from other surveys of mental health among college-aged students, the mental health indicators on the revised ACHA-NCHA will demonstrate evidence of criterion-related validity in the pre-test sample	This hypothesis is partially supported. Past 12-month prevalence estimates are closest between the ACHA-NCHA and NCS-R for depression, Obsessive Compulsive Disorder, panic attacks, and Attention Deficit and Hyperactivity Disorder.* Estimates were less comparable for Anxiety, Bipolar Disorder, Bulimia, Anorexia, and Phobia.
• **Hypothesis 4B:** As a result of analyses demonstrating expected theoretical relationships between variables, mental health indicators on the revised ACHA-NCHA will demonstrate evidence of construct-related validity in the pre-test sample.	This hypothesis is supported. Stress, difficult life events, and recency of negative affect all appear to demonstrate evidence of convergent and discriminant validity. Ordinary least squares and logistic regression models demonstrate relationships between constructs that are consistent with literature and theory.
Research Question 5. *How do mental health indicators on the revised ACHA-NCHA differ from those on the original ACHA-NCHA in terms of their reliability and validity*	
• **Hypothesis 5:** Mental health indicators on the revised ACHA-NCHA will demonstrate evidence of higher levels of reliability and validity than those on the original ACHA-NCHA.	This hypothesis is supported. The negative affect scale on the revised ACHA-NCHA has a higher degree of internal consistency ($\alpha = .893$) than that on the original ACHA-NCHA ($\alpha = .881$). Additionally, past year prevalence estimates collected from the revised ACHA-NCHA* for anorexia, anxiety, bulimia, depression, and substance abuse are closer to the NCS-R estimates than those collected by the original ACHA-NCHA.

*These are the conditions as defined on the ACHA-NCHA. Many conditions had different names on the NCS-R as explained in Chapters 4 and 5.

Summary and Discussion of Results

This study asked five research questions and tested six hypotheses, all of which are delineated in both Chapters One and Three. All hypotheses were at least partially supported by the study results. The following is an overview and discussion of the findings for each research question and its associated hypotheses.

Research questions 1 and 1A

Nearly all of the mental health indicators on the revised ACHA-NCHA had never been used to collect data from the target population of U.S. college students previously. Subsequently, as is the case with most survey pre-tests (e.g., Singleton & Straits, 1999; Fowler, 2004; Tourangeau, 2004), it was first of interest to determine what patterns of student responses emerged in the data collected from the 44 items measuring student mental health in this sample of 6,216 students from 7 U.S. PSIs. Results from all 44 items were examined for the total sample and by gender. No hypotheses were tested for this research question, but Research Question 1A did ask about how student responses varied by demographic characteristics such as gender, year in school, and race/ethnicity. Several interesting patterns emerged in the collected data, many of which are discussed below. However, because no hypotheses were tested and very few results for this research question were analyzed for statistical significance, it is important for the reader to understand that few inferences can be made from the results from Research Questions 1 and 1A.

It should first be noted that students responded at high frequencies to each of the mental health questions on the revised ACHA-NCHA survey. No one item or series of items appeared to have a significant problem with student non-response. The literature

(Fowler, 2004; Aday & Cornelius, 2006) and the Model of Survey Response

(Tourangeau, 1984; Tourangeau et al., 2000) both suggest that if a particular item

demonstrates an especially high degree of non-response, the question may be problematic

for respondents. Hence, this lack of differential non-response provides preliminary

evidence that no single item or series of items was initially problematic in terms of

respondents' comprehension or desirability to respond (Fowler, 1995, 2004; Aday &

Cornelius).

Chapter Four first presents patterns of responses to the 8 items on the ACHA-

NCHA negative affect scale. Results indicate that there is substantial variability in the

level to which students report each of these indicators of negative affect, which is a

desired characteristic of survey questions (Aday & Cornelius, 2006). The magnitude with

which students reported each indicator of negative affect in the last 2 weeks was as

follows (in decreasing order of endorsement): felt overwhelmed by all you had to do

(45.5 percent); felt exhausted—not from physical activity (42.4 percent), felt very lonely

(21.9 percent), felt very sad (21.9 percent), felt overwhelming anxiety (14.5 percent) felt

things were hopeless (14.4 percent), felt overwhelming anger (9.8 percent), and felt so

depressed it was difficult to function (8.5 percent).

When responses are collapsed to past 30 day and past 12 month prevalence,

estimates are not surprisingly much higher for each of these feelings. For example, 15.0

percent of students have felt so depressed it was difficult to function in the last 30 days,

and 31.7 percent have felt this way in the last 12 months. When looking across a larger

time frame, it is alarming to see the frequencies with which students report these feelings.

As is suggested in the literature (e.g., Sharkin, 2006; Kadison & DiGeronimo, 2004;

Kadison, 2006; Cook, 2007; Ross et al., 1999), these data support the fact there appears to be a critical need to assist students in effectively managing these negative emotions in order to prevent the often devastating intrapersonal and interpersonal consequences associated with a negative affect.

In addition to examining estimates collected from the overall sample, the negative affect indicators were also individually examined by gender. Women appear to endorse each of these indicators with higher recency than men. These gender-specific patterns, though not evaluated for statistical significance, are supported by literature in the field that indicates women are more likely than men to report experiencing negative emotions (Kadison & DiGeronimo, 2004) and to have a generalized negative affect (Joiner & Blalock, 1995). It should be noted, however, that these results do not necessarily mean that males *experience* these emotions less frequently than females; as suggested by Kadison & DiGeronimo, females "…are more tuned in to their feeling and tend to see the connection between the stresses of college life and their feelings" (p. 32). The smallest observed difference between gender was for feeling overwhelming anger: 9.1 percent of males and 10.3 percent of females reported experiencing overwhelming anger in the last 2 weeks. It is of great interest that anger, an emotion historically associated with men and masculinity (e.g., Azar, 2007) was reported with similar frequency among women in this sample of college students. The largest observed difference was for feeling overwhelmed by all they had to do: 35.5 percent of males and 52.2 percent of females report this emotion in the last 2 weeks. These data suggest males and females are *reporting* differential experiences with these emotions, an idea that is certainly supported by the literature in the field. Therefore, observed patterns by gender should be further explored

as they may lend additional merit to the validity of the data collected from this revised

survey for indicators on the negative affect scale.

The ACHA-NCHA also collected information on past 12-month prevalence (i.e.,

diagnosis and/or treatment) of 15 mental health conditions, which is a large increase from

the 5 condition estimates on the original survey. As shown in Table 4.2, the prevalence

estimates varied tremendously in this sample with this highest prevalence estimates for

depression (9.3 percent) and anxiety (8.5 percent) in the last 12 months, and the lowest

for non-substance related addictions (0.5 percent) and schizophrenia (0.2 percent). These

prevalence estimates suggest those mental health disorders (i.e., anxiety disorders and

mood disorders) most commonly experienced by the U.S. population as a whole, anxiety

and depression—see Kessler et al., 2005—are also those that are most commonly

experienced by U.S. college students in this sample. Again, patterns observed by gender,

though not evaluated for statistical significance, appear to be consistent with those found

in the literature for both college and non-college populations. For example, women in this

sample report much higher past 12 month prevalence of both depression and anxiety, a

phenomena that has been well established in the general U.S. population (Kessler et al.,

2005; NIMH, 2007b, 2007c). Additionally, men in this sample report substance

abuse/addiction, other addictions, and ADHD with higher frequencies than women in the

last 12 months. The literature further supports these patterns (Presley et al., 1998; Kessler

et al., 2005).

As a whole, 17.1 percent of students in this sample reported being diagnosed

and/or treated with at least one mental health condition in the last 12 months, which is

less than the past-year prevalence estimate (26.2 percent) of any mental health condition

in the general U.S. population (NIMH, 2007a). Of those students in this sample reporting

experience with a mental illness, more than half (55.0 percent) reported diagnosis and/or

treatment for more than one condition on the list of 15 provided. The most commonly

reported co-morbidities were of depression and anxiety, which are well known to co-

occur (NIMH, 2007c). The extent with which students reported co-morbidities is both

consistent with the literature (45 percent; NIMH, 2007a) and alarming in terms of public

health prevention efforts, which will be discussed later in this Chapter.

Table 4.2 additionally presents information on students' reported experiences with

difficulties in the last 12 months. Ten categories were presented to students and the

revised survey asked if students had found any of them to be "difficult to handle" within

the last 12 months. Women reported experiencing significantly more difficulties (mean:

3.45) than men (mean: 2.47) in the last 12 months. As mentioned above, however,

because these are self-report data results do not necessarily indicated that females are

experiencing these difficulties at higher rates than males; rather, they are *reporting* them

with higher frequency than males, which may be a result of their emotional

connectedness and maturity when compared to college males (Kadison & DiGeronimo,

2004).

The top three difficulties experienced by students in this sample—academics,

non-familial relationships, and finances—were consistent across gender. Because

experience with difficult life events is highly correlated to stress and negative affect (see

results for Research Question 4 as well as Ross et al., 1999; Grace, 1997; Cook, 2007),

these results suggest public health practitioners may want to especially focus health

promotion and health education efforts on teaching students how to cope with difficulties

associated with academic, relationship, and financial struggles. Again, findings from this study support recent literature in the field. Sharkin (2006) states,

> "Some counselors believe that the pressure to succeed academically has never been higher…Once accepted, students feel pressure to excel and show that they are worthy of having been accepted. Because the cost of college has skyrocketed over the past several years, students feel intense pressure to achieve academic success to justify their financial commitment…" (pp. 8-9).

Kadison and DiGeronimo (2004) support these suggestions and cite academic demands, financial concerns, and relationship (particularly coupled relationship) difficulties as top contributors to student distress today.

An important new area of interest on the revised ACHA-NCHA was that of help-seeking. Help-seeking is a protective health behavior, and students who seek help for mental health distress are more likely to experience positive outcomes, reduce their stress levels, and complete their college degree (Kadison & DiGeronimo, 2004). Accordingly, the ACHA was especially interested in understanding those student characteristics associated with seeking help. Six new indicators were added to the revised ACHA-NCHA to ask students about help-seeking: four asked about whether they had ever received services from specific mental health providers, one asked about if they had received services from their current university's Counseling or Health Services, and one asked about their willingness to consider seeking help in the future.

Several interesting patterns that warrant further exploration and analyses were initially observed in these data. First, more than one-third of students (36.4 percent) reported receiving mental health services from at least one type of provider (a counselor or psychologist, psychiatrist, other medical provider, and/or a clergy member) for their mental health needs in their lifetime. In this sample, students most commonly reported utilizing counselors/psychologists (30.8 percent) and least commonly reported utilizing

clergy (6.2 percent) for their mental health needs. Thirteen percent of students reported

that they had received mental health services from their current college or university, and

a promising 67.8 percent of students reported they would be willing to consider seeking

help from a mental health professional in the future if something was really bother them.

Women report both utilizing these providers and a willingness to seek help in the future

with higher frequencies than men in this sample, though these differences are not

evaluated for their statistical significance. However, these patterns are consistent with

literature that suggests women are more likely to seek help for mental health distress

(Kadison & DiGeronimo, 2004), thereby adding further preliminary evidence of validity

of these data. Again, this does not necessarily suggest women have greater mental illness

than males; as specified by Kadison & DiGeronimo, "Rather than explore [negative]

feelings (as females will do), [males] avoid even thinking about their feelings and turn

instead to counterproductive coping behaviors—commonly alcohol and drug use" (p. 32).

In addition to help seeking, self injury was a new construct measured by the

revised ACHA-NCHA. A total of 5.6 percent of students in this sample (5.8 percent of

women and 4.0 percent of men) reported intentionally injuring themselves within the last

12 months, which is slightly lower than the 12-month prevalence estimate of 7.3 percent

cited by Whitlock and colleagues (2006). As indicated in Chapter Two, this problem

appears to be on the rise on college campuses and is one that certainly seems to warrant

further study. Furthermore, a comparable number of students in this sample (6.0 percent)

reported seriously considering suicide in the last 12 months, and 0.8 percent (49 students)

reported attempting suicide in the last 12 months. Relationships between self-injury,

suicide ideation, and suicide attempts were beyond the scope of this study; however, the

prevalence of both self-injury and suicide ideation in this sample suggests further research is both warranted and necessary to better understand the characteristics of students who engage in these behaviors to prevent their associated devastating consequences.

In addition to examining all 44 indicators by gender, 13 indicators were further examined by race/ethnicity and year in school. Several interesting patterns that justify further exploration and research were observed in these data. Again, none of these patterns were examined for statistical significance, so no inferences can be made. Furthermore, many of these observations were based on very small sample sizes and are hence likely to have large standard errors associated with their measurement.

Initial results collected from the ACHA-NCHA mental health indicators suggest that Black and Hispanic/Latino students report past 12 month and lifetime prevalence of depression at similar rates to one another and at rates that are lower than White students. These students also appear to seek mental health treatment from clergy more frequently than White students, a phenomena that is consistent with results from previous studies (Sharkin, 2006). Asian students report the lowest frequencies of any racial/ethnic identities of diagnosis/treatment of depression and anxiety in the last 12 months, ever being diagnosed with depression, and all forms of help-seeking in this sample. Again, these patterns are comparable to those presented in the literature. Sharkin (2006) suggests that students of color are more likely to seek mental health services not directly linked with counseling centers; instead, they may prefer academic advisors, financial aid centers, multicultural affairs officers, and religious and spiritual advisors (p. 90). Furthermore, when students of color do seek mental health services on campus, they are

more likely than White students to discontinue them prematurely (Brinson & Kottler, 1995, as cited in Sharkin, 2006).

In terms of year in school, interesting patterns appear to emerge between 2nd and 3rd year undergraduates. Third-year undergraduates report experiencing diagnosis and/or treatment with anxiety and depression (9.2 percent and 11.0 percent, respectively) with higher rates than second-year undergraduates (6.5 percent and 6.6 percent, respectively). This increase in prevalence estimates appears higher than that observed between 1st- and 2nd-year undergraduates and 3rd- and 4th-year undergraduates. Furthermore, 5th-year undergraduates also appear to report past 12 month experience with anxiety (14.6 percent) and depression (15.4 percent) at rates much higher than other students in this sample, and nearly one-quarter (24.9 percent) of these students have been diagnosed with depression in their lifetimes. Three in five 5th-year or more undergraduates report "more than average" or "tremendous" stress in this sample, compared to 50.1 percent of the overall sample. Hence, these patterns seem to warrant further exploration into students' development throughout their time in college and may suggest a need for increased prevention and education targeted between the transition from underclassman to upperclassman status as well as for those students who take longer than the traditional four years to complete their undergraduate degrees. Additional research is necessary to better understand if (1) these differences are statistically significant across year in school and (2) the underlying risk factors associated with these transitions that may put students at higher risk for mental health distress.

In summary, an almost infinite number of intriguing patterns emerged on the 44 revised mental health indicators. The previous summary was in no means exhaustive, but

it does highlight some of the interesting responses that may justify future exploration. Many observed patterns, particularly those that emerged by gender, coincide with those found in the literature for both the U.S. population as a whole and for college students in particular. In addition, no item appears to have a significant problem with student non-response or a pattern of responses that elicits tremendous concern in terms of its psychometric properties. Therefore, the emergence of these patterns enhances support that these data collected from the revised ACHA-NCHA mental health indicators demonstrate evidence of validity, a concept that is further explored throughout this chapter.

Research question 2 and hypothesis 2

Because the original ACHA-NCHA had 13 indicators (5 indicators of negative affect, suicide ideation, suicide attempts, 5 past year prevalence estimates, and ever diagnosis with depression) of students' mental health status that were comparable to revised ACHA-NCHA indicators, it was of tremendous interest to determine how changing the way these questions were asked changed patterns in student responses across versions of the survey. It was hypothesized that changes to survey questions measuring similar constructs across versions would result in significant differences in patterns of student responses. As specified in Chapter Four, this hypothesis was partially supported by the data collected from the two versions of the survey.

Because the sample sizes across both versions of the survey were very large—more than 6,000 respondents in each—there was sufficient power to detect even very small effects. Hence, only those comparisons that resulted in Phi coefficients of 0.1 or higher (a standard for small effects; Miles & Gilbert, 2005) were considered practically

significant and in support of the hypothesis in this study. Although 12 of the 13 direct comparisons indicated statistically significant differences in responses across versions of the survey (see Tables 4.7 and 4.8), only 4 of the 13 comparisons made resulted in differences in student responses with phi coefficients of 0.1 or higher. All four of the comparisons that resulted in practically significant differences were for indicators of negative affect: felt very sad ($\Phi = -.135$), felt things were hopeless ($\Phi = -.138$), felt exhausted—not from physical activity ($\Phi = -.121$), and felt so depressed it was difficult to function ($\Phi = -.116$). No practically significant differences were observed for feeling overwhelmed in the last year, for past year prevalence estimates for anorexia, anxiety, bulimia, and substance abuse/addiction, for past year and lifetime prevalence of depression, or for past year suicide ideation and attempts.

There are multiple potential reasons suggested by the literature and by theory as to why practically significant differences were observed for four of the five comparable indicators of negative affect. Primarily, there were multiple and considerable differences between versions of the survey in the ways these indicators were measured. First, the question stem on the original version of the survey read, "Within the last school year, how many times have you:" with the response options, "Never, 1-2 times, 3-4 times, 5-6 times, 7-8 times, 9-10 times, and 11 or more times" while on the revised version the stem read, "Have you ever:" with the response options "Never, Not in the last 12 months, In the last 12 months, In the last 30 days, In the last 2 weeks." Subsequently, the time frame over which students were asked to respond, the question stem, the response option format, and the number of response options were all different across versions of the survey. Unfortunately, it is not possible to determine the impact of each of these changes

individually, but the combination of these multiple modifications appears to significantly

and practically impact the process of student response when describing their feelings over

the past year in comparable samples.

The response options for the negative affect indicators were modified across

versions because feedback from student participants indicated that they had a very

difficult time remembering exactly how many times they had felt these emotions within

the last school year. The time frame was changed to "In the last 12 months" because the

phrase "Within the last school year" could indicate anywhere from a 4 to 9 month time

period on the original survey depending on when each PSI surveyed and when the fall

semester began. Therefore, the ACHA revisions committee believed they would be able

to collect more precise information from students using the revised stem and response

options. It was believed that the specific time periods would assist students in more

accurately retrieving memories of experiences with these emotions, as specific time

constraints consistent with behavior have been shown to assist respondents with more

accurate recall (Di Iorio, 2005; Loftus, Smith, Klinger, & Fielder, 1994). In addition, the

literature not surprisingly suggests that more recent events are easier for people to recall

when responding to questions (Fowler, 1995; Loftus et al., 1994). When response options

were collapsed to make the items as comparable as possible, more students reported

experiencing these emotions in the last year on the original survey than on the revised

survey. This could perhaps be because students were more likely to employ end aversion

strategies (Di Iorio, 2005, see further discussion below in the discussion of study

limitations) when judging which response option to select on the original survey because

of their lack of precise recall. The one indicator of negative affect that yielded

statistically significant yet practically insignificant (according to the standards specified

in Chapter Three) differences was for feeling overwhelmed—not from physical

activity—within the last year (Φ = -.098). On both versions of the survey, an

overwhelming majority of students (original: 92.9 percent, revised: 87.1 percent) reported

this feeling within the past year. Hence, the variability observed in the other indicators

across versions may not have observed for this indicator because students so commonly

reported this feeling.

Comparisons were also made for past year prevalence estimates of five mental

health conditions across versions of the survey. Interestingly, although there were

statistically significant differences observed in all of the past year prevalence estimates

for mental health conditions across versions of the survey (i.e., anorexia, anxiety,

bulimia, depression, and substance abuse), none were practically significant according to

the specifications set forth in the study methods description. Phi coefficients for the chi-

square analyses ranged from -.022 to -.098, and all prevalence estimates were lower for

the revised version of the survey. The ACHA revisions committee changed the survey

from asking students if they "had" each specific condition within the last school year on

the original version to asking students if they had been "diagnosed or treated by a

professional with each condition" in the last 12 months on the revised version. This

change was made in order to refine the measurement of the constructs and reduce error in

response and interpretation. A variety of self-assessments are available for each of these

concepts, and by changing the stem and response options across versions of the survey to

be more clear and specific (i.e., diagnosed or treated by a professional), the ACHA-

NCHA revisions committee was attempting to ensure students were endorsing these

items only if they had been diagnosed or treated by a health professional for these

conditions. Hence, the committee expected the prevalence estimates to decrease between

versions of the survey because students would not be including self-diagnosis. This

anticipated pattern did emerge for each of the five comparisons of past-year prevalence

estimates, which provides evidence of face validity. Though the effect sizes of each of

these differences were relatively small (less than 0.1), larger differences in effects were

found for more common conditions (anxiety, depression, and substance abuse) than less

common conditions (anorexia and bulimia), which the committee also expected.

The difference in lifetime prevalence of depression was also evaluated across

versions of the survey. The original version of the survey asked students, "Have you ever

been diagnosed with depression?" and the revised version asked students, "Have you ever

been diagnosed (by a professional) with depression?" Again, the survey revisions

committee made this modification to increase precision in measurement because of the

availability of self-assessment screening tools. As was the case in past-year prevalence

estimates, it was expected that the lifetime prevalence estimate—if it changed at all—

would likely decrease as self-diagnoses would be eliminated from the measurement.

Results indicate a statistically significant difference in lifetime prevalence estimates

across versions of the survey, though the effect size of this difference is very small (Φ =

.022) and has little practical significance. Furthermore, a higher percentage of students

reported lifetime diagnosis with depression on the revised survey (14.9 percent) than on

the original survey (13.0 percent), which differs from what was expected by the

committee and by this researcher. A variety of phenomena could be behind these

observed differences (e.g., students may feel more comfortable self-reporting a lifetime

diagnosis with depression on a survey with 44 mental health indicators than one with 18),

but because the observed difference is of such small practical significance these

phenomena will not be further discussed.

Lastly, results associated with this hypothesis also found that little differences in

student responses patterns existed for suicide ideation and attempts across versions of the

survey. Students reported considering suicide in the last year with more frequency on the

original version of the survey (8.4 percent) than on the revised survey (6.0 percent), and

although results were statistically significant ($\chi^2 = 26.385$, df = 1, $\Phi = -.046$), the effect

size of the differences again bears little practical significance. There were no statistically

significant difference in the proportion of students reporting suicide attempts in the last

year across versions of the survey, which is not surprising given that a suicide attempt is

likely a very memorable event in a student's life, and as such may not be as prone to

various types of measurement error in terms of the comprehension, retrieval, and

judgment stages of the Model of Survey Response (Tourangeau, 1984; Tourangeau et al.,

2000; Loftus et al., 1994). Furthermore, suicide attempts are reported by only a very

small proportion of students in both versions of the survey, and a more focused as a

opposed to general study among students who have experienced mental health distress

may be warranted to determine if differences do exist in the ways students respond to this

question across versions of the survey.

Ultimately, differences in survey questions across versions of the ACHA-NCHA

did produce statistically significant differences in the patterns of student responses. For

all indicators except lifetime diagnosis with depression, students reported experiences

with lower frequencies on the revised version than on the original version of the survey.

Because the committee worked to deliberately refine the measurement of student

responses, this decrease in endorsement across versions is not at all surprising—in fact, it

provides evidence that these measurements may be more precise. The effect sizes of these

differences vary; only four met the cutoff of standard "small" effects, and all four were

for indicators of negative affect. Because feelings are latent and transient in nature and

accordingly are not easily quantifiable, they may be more subject to recall bias and to

instrumentation impacts (Fowler, 1995). On the other hand, suicide ideation and attempts

and experience with a mental illness may be more specific life events that are more easily

remembered by students, which may be a reason why instrumentation changes did not

make as large a difference across versions of the survey (Fowler, 1995). These observed

changes should be noted by any institution that has previously surveyed its students using

the original ACHA-NCHA and intends to switch to the revised ACHA-NCHA once it is

launched. Institutions must be made aware that differences may exist solely because of

the way the construct is measured—and not because of campus-wide prevention or

intervention effects. These implications are discussed further later in this chapter.

Research question 3 and hypotheses 3A and 3B

The third research question and associated hypotheses in this study evaluated the

extent to which the negative affect scale on the revised ACHA-NCHA was reliable.

Results of the principal component analysis indicate this scale is unidimensional in

nature, and each of the eight variables on the scale (felt things were hopeless, felt

overwhelmed by all you had to do, felt exhausted—not from physical activity, felt very

lonely, felt very sad, felt so depressed it was difficult to function, felt overwhelming

anxiety, felt overwhelming anger) all appear to be indicators of the same underlying

construct (i.e., a negative affect). The reliability analysis suggests the eight indicators have a coefficient alpha of .893 and that alpha does not increase if any indicators are removed from the scale. This coefficient alpha is well above the standard cutoff of .700 specified by Nunnally and Bernstein (1994), suggesting excellent internal consistency of measurement. Therefore, both Hypothesis 3A and 3B were both fully supported by these data.

As is discussed in Chapter Two, reliability is a necessary but insufficient condition for validity of measurement. Hence, because these indicators appear to demonstrate sufficient evidence of reliability, they may be further examined for evidence of their validity. The correlation matrix associated with the indicators of negative affect provides evidence of construct-related validity, as correlation coefficients are often used to simultaneously examine both convergent and discriminant validity (Spector, 1992). Because it is expected that each of these variables is a unique indicator of negative affect, all were anticipated to correlate with one another (thereby providing evidence of convergent validity), which was observed. Correlation coefficients ranged from a low of .328 for the relationship between "felt overwhelmed by all you had to do" and "felt overwhelming anger" to a high of .789 for the relationship between "felt very lonely" and "felt very sad". As none of the coefficients are .800 or higher—a commonly accepted cutoff for multicollinearity (Hensher, Rose, & Greene, 2005)—each indicator appears to be measuring an independent concept (thereby providing evidence of discriminant validity). This combination of convergent and discriminant evidence of validity, as well as the fact that alpha associated with the negative affect scale does not increase if any indicator is removed suggests each indicator contributes a uniquely important amount of

information to the scale. Additional evidence of construct-related validity of the negative affect scale is discussed below in the summary and discussion of Research Question 4 and Hypothesis 4B.

Research question 4 and hypotheses 4A and 4B

The fourth research question and associated hypotheses were concerned with evidence of validity of the data collected from the revised ACHA-NCHA mental health indicators. Two aspects of validity—criterion-related and construct-related—were examined in order to attempt to ensure this survey was in fact measuring what it purported to measure.

Table 4.12 present a comparison of 12-month prevalence estimates collected from the ACHA-NCHA and the National Co-morbidity Study Replication (NCS-R). As discussed in Chapter Four, results indicate that the prevalence estimates generated by the revised ACHA-NCHA vary greatly in the extent to which they compare to those generated by the NCS-R. Relatively comparable estimates appear to be present for depression (9.3 percent) in the ACHA-NCHA sample compared to "any mood disorder" in the NCS-R (9.7 percent), Obsessive Compulsive Disorder in both samples (ACHA-NCHA: 1.6 percent, NCS-R: 1.2 percent), Panic Attacks on the ACHA-NCHA (4.0 percent) compared to Panic Disorder in the NCS-R (2.7 percent), and Attention Deficit and Hyperactivity Disorder in both samples (ACHA-NCHA: 2.9 percent, NCS-R: 4.1 percent). Estimates that are limited in their comparability include 12-month prevalence of Anxiety on the revised ACHA-NCHA (8.5 percent) compared to Generalized Anxiety Disorder from the NCS-R (2.7 percent), Bipolar Disorder (ACHA-NCHA: 1.1 percent, NCS-R: 2.8 percent), Bulimia (ACHA-NCHA: 1.1 percent, NCS-R: 0.3 percent), and

Anorexia (ACHA-NCHA: 0.9 percent, NCS-R: 0.0 percent). Perhaps the biggest

disparity between surveys is the estimate for Phobia on the ACHA-NCHA (0.9 percent)

and Specific Phobia on the NCS-R (9.1 percent). Many of the differences observed

between the NCS-R and the revised ACHA-NCHA might be expected based on findings

in the literature. For example, college students have been known to be at especially high

risk for eating disorders (i.e., anorexia and bulimia) compared to the general population

(Phillips & Pratt, 2005; Schwitzer et al., 1998; Drewnowski et al., 1994). However,

further research is needed to better understand the complexities associated with

differences in the results between survey instruments, which are likely due to differences

in the ways constructs are operationalized and measured.

It should be noted that many similar patterns existed by gender on the revised

ACHA-NCHA and NCS-R. Comparable patterns across gender (e.g., women have higher

prevalence rates than men or vice versa) were observed for anxiety, ADHD, bipolar

disorder, bulimia, depression, OCD, panic attacks, and substance abuse; patterns differed

across gender for anorexia and phobias. Very few patterns were similar between age

groups: only for anxiety and panic attacks did similar trends between younger adults and

older adults (i.e., younger adults reported lower prevalence estimates than older adults)

appear on both the ACHA-NCHA and the NCS-R. For all other estimates (anorexia,

ADHD, bipolar disorder, depression, OCD, phobias, and substance abuse), the patterns

between 18-29 year olds and 30-44 year-olds differed between ACHA-NCHA and NCS-

R results. Because only a small proportion (5.3 percent) of the revised ACHA-NCHA

sample was aged 30-44 years, these results must be interpreted cautiously. Further

research is necessary to determine if there are true differences between prevalence

estimates across age groups on the revised ACHA-NCHA and the extent to which non-traditionally aged students differ from non-student adults of comparable ages.

Because many estimates appear to differ somewhat substantially between the revised ACHA-NCHA and the "gold standard" prevalence estimates collected by the NCS-R, Hypothesis 4A was only partially supported by these results. There are several limitations of these particular comparisons that must be addressed independently of the overall study limitations discussed later in this chapter. First, the NCS-R generalizes to the U.S. resident population as opposed to solely college students. Individuals who attend colleges and universities are known to differ from those who do not. Therefore, the populations are not directly comparable and results must be interpreted with caution. Second, the NCS-R utilizes a structured in-person interview format as compared to the online format of the ACHA-NCHA. Responses to online surveys have been shown to differ from other forms of surveys in the literature (Daley et al., 2003); hence, even if participants were directly comparable, some differences in measurement would likely be expected due to survey modality. Third, and perhaps most importantly, the ACHA-NCHA and the revised NCS-R conceptualize the prevalence of a variety of mental health constructs very differently. This difference in construct conceptualization make direct comparisons impossible for a variety of mental health conditions. Nonetheless, because these are best population prevalence estimates available, it is still useful to make comparison to begin to evaluate evidence of criterion-related validity—particularly across versions of the survey, as described below in the summary of Research Question and Hypothesis 5. Readers are encouraged to understand, however, that all comparisons are purely qualitative. In other words, though numerical data are presented, no statistical

comparisons have been made and as such no direct inferences can be made about the true extent to which these estimates differ from the revised ACHA-NCHA to the NCS-R and as such the level to which criterion-related validity can truly be established is marginal.

Hypothesis 4B suggested that data collected from the revised ACHA-NCHA would demonstrate evidence of construct-related validity; in other words, statistical analyses would yield relationships expected between constructs as they are delineated by both theory and previous studies. This hypothesis was fully supported by the data collected from the ACHA-NCHA. All three models designed to test this hypothesis produced statistically significant and meaningful results, and the examination of R^2 values suggested that appropriate levels of variance in outcomes was explained by the predictors of interest in the models, thereby providing evidence of both convergent and discriminant validity.

The first model in the study predicted level of negative affect (a standardized factor score derived from the 8 indicator variables of the negative affect scale) from self-reported stress levels while controlling for race/ethnicity, year in school, age, and gender. When only demographic characteristics were used to predict students' level of negative affect, the model accounted for only 5% of the variance in the outcome. All demographic variables were statistically significant predictors of negative affect, though because of the large sample size there was enough power to detect even very small differences. When assuming no model misspecification and holding all else constant, gender appears to be the most relatively important predictor of negative affect (b = .396, β = .196, p < .001). After controlling for all other demographic variables, college women report a recency of negative affect that is .396 standard deviations higher than college men, which is

consistent in many ways with findings from Research Question 1. Racial/ethnic differences also exist in recnecy of negative affect: White students have higher recency of negative affect when compared to a collapsed group of Black, Hispanic/Latino, and Asian students (b = -.093, p < .01), and lower recency of negative affect when compared to a collapsed group of Bi/Multiracial and Other racial identity students (b = .241, p < .001).

When stress was added as a predictor to the model, all of the predictors of interest accounted for 27.1 percent of the variance in negative affect. Year in school was no longer a statistically significant predictor, and the relationship between Black, Hispanic/Latino, and Asian student grouping and White students was no longer statistically significant once stress was added to the model. This indicates that the relationships observed in the model without stress as a predictor may be mediated by students' stress levels. In the final model, stress is the most relatively important predictor of negative affect (b = .589, β = .484, p < .001), and as students' response options changed by one response category (e.g., average stress to more than average stress), their recency of negative affect level changed by one .589 standard deviations. Gender remained the most significant demographic predictor in the model; when controlling for stress and all other demographic variables, women reported a recency of negative affect that was .237 standard deviations higher than men. Interestingly, when stress levels were included in the model, the relationship between gender and negative affect was reduced, thereby indicating that it is likely partially mediated by stress levels. All of these findings are consistent with both theory and literature (e.g., Kadison & DiGeronimo, 2004). Therefore, findings suggest that the stress indicator and recency of negative affect scale (comprised of 8 indicators) show evidence of construct-related validity.

The second model in the study predicted students' self-reported stress levels from various demographic characteristics and from the top 5 most commonly experienced difficulties (i.e., academics, career-related issues, non-familial relationships, finances, and personal appearance) in this sample. Results indicate that demographics alone accounted for little variability in the outcome (R^2 = .053). Once again, all demographic variables were statistically significant predictors of students' stress levels when controlling for one another, and gender was the most relatively important predictor (β = .166) followed by year in school (β = .114) As was the case in the previous model, White students reported significantly higher stress levels than the combined group of Black, Hispanic/Latino, and Asian students, and significantly lower stress levels than Bi/Multiracial/Other racial identity students.

When the top five most commonly experienced difficulties were added into the model, all of the predictors of interest accounted for nearly one-fifth of the variance (R^2 = .196) in students' stress level. Each difficulty was entered into the model as a dummy variable with "No" responses as the reference group. Accordingly, positive b and β values indicate that stress levels increased as students reported the difficulty, which is the relationship one would expect based on previous literature and theory. When controlling for all other difficulties and demographic traits, academic difficulty is the most relatively important predictor in the model (β = .236), followed by gender (β = .103), difficulty with non-familial relationships (β = .099), and year in school (β = .098). Again, as was the case in the first model, the observed difference between White students and Black, Hispanic/Latino, and Asian students became statistically insignificant, indicating the relationship between race/ethnicity and self-reported stress level may be mediated by

experiences with difficult life events. Ultimately, this model suggests that all five of the most commonly experienced life events are significant predictors of students' stress levels when controlling for one another and for race, age, year in school, and gender. This suggests that they each are independent constructs (i.e., possess evidence of discriminant validity). Furthermore, this model indicates that colleges and universities may reduce both stress and negative affect (because, as shown in the first model, stress predicts negative affect) by helping students effectively cope with difficulties, particularly those associated with academic demands and non-familial (i.e., partnerships, friendships) relationships.

The third and final model tested to examine evidence of construct-related validity predicted past year diagnosis and/or treatment with depression from demographic characteristics, stress level, and recency of negative affect (once again a standardized factor score created from the 8 indicator variables of the negative affect scale). Results of the first model tested, that which evaluated only demographic predictors, suggest that females were more than twice (OR = 2.276, p<.001) as likely as males to report diagnosis and/or treatment with depression when controlling for all demographic predictors. As was the case in the first two linear regression models, the combined grouping of Black, Hispanic/Latino, and Asian students appeared to be at lower risk (OR for depression = .530, p<.001) than White students. However, no statistically significant differences were observed in risk for diagnosis/treatment with depression between White and Bi/Multiracial/Other students when controlling for all other demographic variables in the first model. Both age (a one year increase produced a .027 unit increase in odds for past-year depression, p <.01) and year in school (a one year increase produced a .115 unit

increase in odds for past-year depression, $p < .001$) were also statistically significant predictors of diagnosis and/or treatment with depression in the last 12 months when controlling for all other demographic variables.

When both level of stress and recency of negative affect were included in the logistic regression model, the odds ratios for each statistically significant demographic predictor decreased slightly. This indicated the relationships between the demographic predictors and the logit of past-year experience with depression may be partially mediated by level of stress and recency of negative affect. When controlling for all demographic variables and for recency of negative affect, a one unit increase in students' stress levels produced a .270 unit increase in the odds of experiencing depression in the last 12 months. Further, when controlling for all other predictors (i.e., demographics and stress level), each one standard deviation increase in recency of negative affect made students 2.388 times as likely to have been diagnosed or treated with depression in the last 12 months. These results suggest that (1) stress and recency of negative affect both have a uniquely important role in predicting students' likelihood of being diagnosed or treated with depression, (2) recency of negative affect appears to be a more relatively important predictor of depression in the last 12 months than stress, and (3) these relationships exist when controlling for various demographic characteristics.

All three models provide evidence of construct-related validity because relationships between data collected from all variables behave as was expected from reviews of the literature and from theory. In addition, these models suggest certain groups of students were at particularly high risk for mental health distress in this sample. For each model, females were at higher risk than males, Bi/Multiracial/Other racial identity

students were at higher risk than White students, and older students and upperclassmen were at higher risk than younger students and lowerclassmen, respectively. These results provide initial evidence that these demographic groups may particularly benefit from focused prevention and education efforts. Need for intervention and areas of future research are discussed later in this Chapter.

Research question 5 and hypothesis 5

The fifth and final research question associated with this study evaluated the extent to which the revised and original ACHA-NCHA mental health indicators differed in terms of their evidence of reliability and validity. As is the case with most survey revisions endeavors, the intent of the intensive survey revisions process was to create a survey that (1) more adequately addressed current student health needs, (2) responded to the needs of participating PSIs and students, and (3) collected data that were more psychometrically sound and less prone to measurement error than those from the original survey. Because the intention was to build upon the original ACHA-NCHA to create as valuable a survey tool as possible, it was hypothesized that data collected from mental health indicators on the revised survey would demonstrate more evidence of reliability and validity than those data collected from mental health indicators on the original ACHA-NCHA. This hypothesis was supported by the results of this study.

The negative affect scale on the revised ACHA-NCHA added three additional indicator variables (felt very lonely, felt overwhelming anxiety, felt overwhelming anger) to the five that were on the original ACHA-NCHA. Further, as described previously, the question stems, response options, recall time period, and format all differed across versions of the survey. Results from the study indicate that internal consistency of the

negative affect scale on the revised ACHA-NCHA negative affect scale was higher (α = .893) than that on the original ACHA-NCHA negative affect scale (α = .881). These findings, which indicate the scale with more indicator variables is more reliable that that with less, are consistent with measurement theory (DeVellis, 2003).

Furthermore, as provided in Table 4.17, data collected from the revised ACHA-NCHA also appear to generate past year prevalence estimates that are more comparable to NCS-R estimates. The limitations of these comparisons have been provided previously; however, it does appear that prevalence estimates for anorexia, anxiety, bulimia, depression, and substance abuse are all closer for the revised ACHA-NCHA than the original when comparing to the nationally generalizable NCS-R. That said, results from Research Question 2 suggest that the differences observed between the revised and original survey prevalence estimates, though statistically significant, have low practical significance. Accordingly, these results are again more qualitative in nature and further research will be necessary to determine which of the two versions of the survey produce more valid data, particularly if true inferences are to be made. The literature on developing survey questions, however, would suggest that the revised past year prevalence estimate questions are likely to collect more valid estimates as they are more specific and more appropriately define for respondents what they are attempting to measure (Fowler, 1995).

Summary of key findings

This study found that students responded to the revised ACHA-NCHA mental health indicators with interesting and appropriate patterns that are supported by current literature. Though they were not analyzed for statistical significance because it was

beyond the scope of this study, interesting patterns of responses to these items appear to emerge by various demographic characteristics, including gender, year in school, and race/ethnicity. Future studies are needed to explore these relationships in more depth. As highlighted in Table 5.1, when compared to responses patterns on 13 items measuring similar constructs on the original ACHA-NCHA (e.g., indicators of negative affect, past year prevalence estimates), student response patterns on the revised ACHA-NCHA differed significantly for nearly all but only practically for 4 indicators of negative affect. This suggests that perhaps latent constructs are more prone to instrumentation differences than tangible life events like the presence of a particular mental health condition.

The negative affect scale on the revised ACHA-NCHA demonstrated evidence of high internal consistency—and a reliability coefficient higher than that on the original ACHA-NCHA—which suggests the three newly added indicators of negative affect improved the reliability of the scale and hence improved the quality of survey measurement as intended by the survey revisions committee. Furthermore, mental health indicators on the revised ACHA-NCHA demonstrate preliminary evidence of both construct- and criterion-related validity in the pre-test sample, though results are more conclusive for evidence of construct-related validity. Estimates collected from the revised ACHA-NCHA are closer to national prevalence estimates than those collected from the original ACHA-NCHA, thereby lending further support that the ACHA achieved its goal in creating a more psychometrically sound instrument during its extensive revisions process. However, because practically significant differences were not observed in past-year prevalence estimates across versions of the survey, ongoing research is necessary to further support these findings.

Study Implications

The implications of this study in both the fields of college student mental health and survey research are numerous. First, this study provides descriptive data on 44 indicators of mental health for a sample of more than 6,000 students attending 7 U.S. PSIs collected from the revised ACHA-NCHA. These data have never been published, and this is the first time that much of this information has been presented in a sample this large and with information from more than one PSI. Thus, this study provides important and interesting information that can help better frame the issue of college student mental health today and can provide possible directions for prevention and intervention. For example, the fact that 17.1 percent of college students in this sample have been diagnosed with or treated for at least one of the 15 mental health conditions assessed on the revised survey in the last 12 months sheds light that college campuses need adequate staff, support, and programming to address these needs. The shift in prevalence of a variety of mental health conditions and concerns between 2nd and 3rd year undergraduates, as well as among 5th year undergraduates highlights these may be critical times to consider for future research and prevention efforts or to explore using qualitative techniques. Further, data collected from revised survey suggest students are willing to seek help, as more than one-third (36.4 percent) had ever received care from a mental health provider, and 67.8 percent were willing to consider seeking help in the future if something was really bothering them. This combination of prevalence data, demographic data, and help-seeking data all provide college health professionals with a better understanding of what is presently happening with college students in terms of their mental health.

Second, data collected from the revised ACHA-NCHA mental health indicators demonstrate evidence of internal consistency reliability, construct-related validity, and criterion-related validity. As such, results support that the revised ACHA-NCHA is an appropriate and useful measurement tool to assess a variety of mental health constructs in college students. Prior to the development of the revised ACHA-NCHA, the original ACHA-NCHA was the only large scale survey available to measure college students' mental health—and the literature notes deficiencies in the content covered by the original ACHA-NCHA mental health section (Soet & Sevig, 2006). Because the data collected from the revised survey provide evidence that they are psychometrically sound, it is believed that PSIs across the country (particularly when they are similar to those that participated in the experimental field pre-test) will be able to use the revised survey and the expanded mental health indicators to capture a better understanding of their students' mental health status and needs in the future. With the national focus on prevention, detection, and intervention on mental health issues among college students (Voelker, 2007; NMHA & Jed, 2002; SPRC, 2004) the ACHA-NCHA is a unique and important source of data at the institutional and national level. Because the data collected from the revised survey's mental health indicators demonstrate evidence of reliability and validity, campuses may feel relatively confident in the estimates and information provided by the survey in terms of students' feelings, stress levels, self-harm and suicide-related behaviors, help-seeking behaviors, and past year prevalence estimates for numerous mental health conditions.

Third, the ACHA-NCHA measures a wide variety of college student health issues in addition to mental health including health, health education, and safety; alcohol,

tobacco and other drug use; sexual health; and impediments to academic performance. Researchers commonly analyze these data to better understand the theoretical and practical relationships between a variety of health constructs (e.g., the relationship between high-risk drinking and depression, the relationship between demographic characteristics and suicide attempts). As mental health concerns on college campuses continue to rise, the evidence of reliability and validity of the 44 indicators on the revised survey suggests that these indicators are appropriate for use in future analyses with a variety of potential variables. Hence, these data have tremendous implications to assist public health professionals in understanding the risk and protective factors for a variety of concerns and how students' mental health status impacts and is impacted by numerous other constructs.

Fourth, this dissertation examined the preliminary reliability and validity of only the revised mental health indicators. This research may therefore serve as a template for future analyses of the reliability and validity of the other revised sections of the survey as well as for evaluating the reliability and validity of indicators on other large-scale surveys measuring public health constructs.

Fifth, this study shows that even minor changes to survey indicators result in statistically significant differences in student response patterns, and these differences may be highest among latent constructs such as feelings as opposed to experiences with mental health conditions. The ACHA must be sure to clearly and explicitly inform participating PSIs that estimates collected from comparable indicators are not comparable across versions of the surveys. Many institutions participate in survey efforts more than once to monitor trends and/or to evaluate program effects (ACHA, 2007b). Accordingly,

as nearly every indicator showed lower pathology on the revised survey (11 of the 13 comparable indicators resulted in statistically lower differences between surveys), if institutions did not understand how instrumentation can impact student responses, they may incorrectly conclude that the prevalence of a variety of conditions had decreased or that students experienced a variety of feelings (e.g., felt very sad, felt so depressed it was difficult to function) at lower rates if the revised survey was used after the original. Hence, for example, if a campus used the original ACHA-NCHA as a program pre-test assessment, and the revised ACHA-NCHA as a post-test assessment for measuring levels of student depression, it would only be able to conclude that their program made an effect if the Phi was greater than that observed in this study.

Sixth, this study shows that an experimental pre-test methodology has tremendous utility in examining how instrument changes impact responses when measuring similar constructs of interest. Marsh (1982) claims, "Surveys and experiments are the only two methods known to me to test a hypothesis about how the world works" (p. 6). This study, and survey experimental pre-tests in general, use both methods to answer the research questions of interest. Hence, the literature overwhelmingly supports this form of survey pre-testing, and this study adds further evidence that it is an appropriate and highly useful method for assessing instrument differences. This research suggests it is imperative for even the smallest of changes to be monitored when surveys are changed to assess how those changes in question wording, response options, and formatting impact how respondents answer the questions. Therefore, this study supports the most current literature in the field of survey research that highlights experimental pre-testing as an appropriate methodology to use when examining the impact of changes to survey items

across versions of the same survey (e.g., Presser et al., 2004; Presser & Blair, 1994; Moore et al., 2004; Tourangeau, 2004; Fowler, 2004). It also supports literature that suggests studies of this nature are feasible, relatively easy to conduct—particularly with the use of online survey applications—and that they provide an abundance of information in terms of student response patterns. As delineated by Presser and colleagues (2004), "To determine whether [survey] revisions are improvements…there is no substitute for experimental comparisons of the original and revised items" (p. 18). This study utilized such techniques, and accordingly allows for more conclusive inferences to be made about whether the revisions process resulted in measurement improvement across versions of the survey.

Lastly, and perhaps most importantly, the results of this study suggest that the ACHA-NCHA revisions committee succeeded in its goal to improve the already useful ACHA-NCHA to better assess and monitor students' mental health experiences. Data collected from the revised survey demonstrate sufficient reliability and validity independently as well as higher internal consistency and preliminary evidence of criterion-related validity than from the original. Hence, this study supports the idea that large-scale surveys benefit from revision and additions to best understand their target populations when the revisions process follows that which is outlined by the field.

Study Limitations

Despite the study's strengths, there are several limitations impacting the inferences made as well as the generalizability of these results that must be addressed. An overview of such limitations is presented below and is arranged by limitations of survey

research, limitations of study design, limitations of study participants, and limitations of analyses and results.

Limitations of survey research

The limitations of survey research are well-documented in the literature, and it has been suggested that in any particular survey, there may be aspects specific to the population, the subject matter, or the data collection methodology that affect the ability of the questionnaire to perform as intended (DeMaio et al., 1998). Because this study is based on survey data, it is first important to address some of these common and relevant issues. Therefore, limitations associated with online data collection, response rates, and respondent errors and bias will be summarized in the context of this research.

Both versions of the ACHA-NCHA collected responses from students using an online survey application. As described in Chapter Two, online surveys have gained popularity in the last decade for numerous reasons, and it has been suggested that data gathered from web-based surveys have the potential to improve both the effectiveness and efficiency of health research and interventions (Karras & Tufano, 2006). Web-based e-surveys are particularly relevant for collecting self-report data from college students because students are technologically savvy, the surveys can be completed at any time of the day, and because e-surveys they have been shown to reduce social desirability bias because respondents are not answering questions directly to an interviewer (Herrero & Meneses, 2006). Yet, there are several limitations of online survey research. First, web-based surveys create data collection environments that cannot be controlled or monitored (Daley et al., 2003), which has the capacity to introduce error into measurement. Second, online surveys have been shown to have significantly lower response rates than in-person

or paper-and-pencil surveys (Couper, 2007). Response rates of online surveys range tremendously, and selection bias occurs due to the non-representativeness of the Internet population and the self-selection of respondents to participate (Eysenbach & Wyatt, 2002). Online surveys discourage participation among those who do not feel technologically-savvy (e.g., older adults), which may have discouraged older students, students without home Internet access, and non-traditional students from participating in this study. They also discourage those who are concerned with WWW privacy issues (Eysenbach & Wyatt). Because the information collected from respondents was very sensitive in its nature (e.g., sexual behaviors, illicit drug use, mental health concerns), students who are concerned with privacy of their personal information on the WWW may not have completed the survey. Hence, these limitations may impact the overall representativeness of the college population targeted with this study. Online survey applications may also experience challenges with web-browser compatibility, firewalls, incorrect student email addresses, and network errors, all of which may make completion difficult for students, thereby reducing the number of responses (Kypri et al., 2004).

Accordingly, perhaps one of the largest issues associated with survey research— and online survey research in particular—is non-response, which "is a problematic, important source of survey error" (Fowler, 2002, p. 56). The average response rate of both versions of the online ACHA-NCHA was less than 18 percent; though this is consistent with response rates for online surveys throughout the literature (Couper, 2007), it still indicates that more than 80 percent of students invited to participate in data collection efforts did not respond to the surveys. People who respond to surveys are known to differ from those who do not (Eysenbach & Wyatt, 2002). They may be more

interested and vested in the issue being researched (Pealer et al., 2001), they are more likely to be female (Eyesbach & Wyatt), and are more likely to be White or Asian than African American or Hispanic/Latino (Cranford et al., 2008). Furthermore, a recent study found additional reasons for student non-response to online surveys included they were "too busy" (45.7 percent), "not interested" (18.1 percent), and they "forgot to complete the survey" (18.1 percent) (Cranford et al.). These issues of non-response are important to consider in terms of representativeness of samples. For the purposes of this study, this means that any descriptive statistics (e.g., prevalence estimates, estimates of student stress) must be interpreted with caution. This is one of the primary reasons this researcher chose not to make inferences about various student groups based on these data; all results for Research Question 1 are purely exploratory. Because very similar response rates were observed across versions of the survey, and because populations seem nearly identical demographically, this issue is not expected to dramatically impact the results and interpretations made to compare items across versions of the survey.

The last limitation of survey research that is certainly applicable to this study is that of respondent errors and bias. A large body of literature exists in regards to these issues, and numerous "response sets" or biases have been documented (DeMaio et al., 1998; Di Iorio, 2005), each of which impacts the validity of the data collected. Response sets lead to sample estimates that are consistently higher or lower than the true values (Aday & Cornelius, 2006). Common response sets that are likely to be applicable to this research include social desirability, acquiescence and nay-saying, end aversion, and recall bias. The issue of social desirability bias is known to impact respondents' answers to health surveys (e.g., Eysenbach & Wyatt, 2002; Di Iorio). Social desirability bias results

in over-estimates of positive behaviors or attitudes and under-estimates of those

behaviors or attitudes that are not perceived as positive. In the context of this study, if a

student believed that help-seeking was viewed as a positive attribute, he or she may be

more likely to endorse the item. On the other hand, if help-seeking was viewed as a

weakness, he or she may be less likely to endorse the item. Hence, issues of social

desirability are important to address, and future research should examine which behaviors

and attitudes are deemed desirable by students. Acquiescence and nay-saying are

converses of one another; research indicates that some individuals are more likely to

choose "yes" responses while other are more likely to choose "no" responses. For

example, in the case of the ACHA-NCHA, those who are acquiescers may be more likely

to say multiple events had been difficult for them to handle in the past 12 months even if

they had not; nay-sayers may have been more likely to say they were not difficult to

handle even if they were. End-aversion (i.e., central tendency) is a response set that

suggests some people avoid choosing extremes on any scale and tend to choose options

that are closer to the middle (Di Iorio). So, for example, in the original version of the

survey, a student may have said he or she had felt hopeless 1-2 times in the last school

year as opposed to "never" in order to be more central. Because the revised version of the

survey had two responses that categorize students into not experiencing this feeling in the

last year ("never" and "not in the last 12 months"—and hence, "not in the last 12

months" was not at the end of the scale) this issue may partially explain some of the

observed differences between versions. Recall bias is the inability of the respondent to

correctly recall the event or feeling of interest. Fowler (1995) suggests there are three

principles relevant to recall: (1) the more recent the event, the more likely it is to be

recalled; (2) the greater the impact or current salience of the event, the more likely it is to be recalled; and (3) the more consistent an event was with the way the respondent thinks about things, the more likely it is to be recalled. Further, he suggests that steps a researcher can take to reduce response distortion include assuring confidentiality of responses, communicating as clearly as possible the priority of response accuracy, and reducing the role of an interviewer in the data collection process. Because the revised ACHA-NCHA set very specific time frames (e.g., last 2 weeks, last 30 days, last 12 months), because it utilized an online application which reduces the role of the interviewer and also appears to improve perceived confidentiality of responses (Aday & Cornelius, 2006), and because mental health issues are likely to be salient for students, it is hoped these issues of recall, as well as other biases, were minimized on the revised ACHA-NCHA.

In addition to all of the above response sets, common reasons for respondent errors include misunderstanding what information was requested, maturation issues such as fatigue or boredom, lack of interest or time, carelessness or guessing, and failure to follow instructions (Di Iorio). As suggested in Chapter Two, each stage of the Model of Survey Response (i.e. Comprehension, Retrieval, Judgment, and Response) is associated with certain types of respondent errors. The extent to which all of these response sets and potential respondent errors operated on the responses given by students on both the revised and original ACHA-NCHA mental health indicators is unknown. As such, this is a limitation of the study, as it is with most research involving the collection of survey data. Future research could utilize more qualitative methods such as cognitive interviewing and respondent debriefing to better understand these phenomena.

Nonetheless, despite their limitations surveys continue to remain the most utilized research tool in sociological and public health studies (Di Iorio, 2005) and are the preferred approach for collecting data from large numbers of students about their college and health experiences (Pealer et al., 2001; Ouimet et al., 2004). Hence, readers are cautioned to interpret results with prudence and to understand the limitations of this data collection methodology.

Limitations of study design

In addition to the limitations discussed in regards to survey research, there are other limitations associated with this study's design that must be highlighted. Most importantly, many questions across versions of the survey had been changed significantly and all question changes were incorporated into the revised survey simultaneously. Hence, it was impossible to gauge the specific and individual impact of each slight adjustment (e.g., stem re-wording, response options) to student responses. Tourangeau (2004) proposed that the most scientifically rigorous design is the fully-crossed factorial design in which all possible combinations of changes are tested. This methodology has the potential to evaluate the impact of each modification as well as the interaction effects of multiple changes. However, these designs are extremely labor-intensive and may result in a "combinatorial explosion" (Tourangeau, 2004); thus, they may be unrealistic for large-scale efforts such as the ACHA-NCHA that are attempting to examine the effects of a multitude of changes to numerous items.

Another limitation of this design is that students responded to one version of the survey or the other, and accordingly it is not possible to determine individual differences between versions of the survey. Asking respondents to take both versions of the survey so

responses could be matched would have been a great deal to ask of the students; subsequently, students from each PSI were randomized to participate in either the revised or original survey. Though randomization controls for a variety of factors, the study design may be strengthened if the same participants were used and responses were matched across versions.

Lastly, another limitation of this study design is that it is almost entirely quantitative in nature. Fowler (1995) provides three possible steps to assess how well questions we ask meet process standards which include focus group discussions, cognitive interviews, and field pre-tests under realistic conditions. This study utilized an experimental field pre-test to evaluate differences between questions across versions of the same survey, but it did not employ the qualitative techniques of focus groups or cognitive interviews. As highlighted in Chapter Two, these strategies would certainly provide relevant and important information about how students approached responding to indicators on the revised ACHA-NCHA. This knowledge would help identify specific challenges experienced by students as they respond to the revised ACHA-NCHA and may help survey developers reduce bias in responses. Hence, future research using these strategies that will allow for the triangulation of both qualitative and quantitative data are recommended, as discussed later in this chapter.

Limitations of study participants

Perhaps the largest limitation of this study is that associated with the self-selection of both institutions and students for participation. As discussed in Chapter Three, PSIs self-selected to participate in the experimental field pre-test of the revised ACHA-NCHA. The seven institutions that opted to participate in the study are very limited in their

generalizability to all U.S. colleges and universities; all are large, all are research institutions, all are 4-year, and all but one are public. Students who attend these institutions are likely to differ significantly from students who attend smaller, 2-year, private, and/or non-research focused institutions. For example, students who attend community colleges are known to differ significantly from those who attend 4-year institutions, and are far less likely to be studied than traditional college students in general. Pascarella and Terenzini (1998) remark,

> "By the time one gets to community colleges, with their open admissions policies, faculties rewarded essentially for teaching, and their disproportionate numbers of non-resident, part-time, older, non-white, and working class students, they are virtually off the radar screen in terms of public recognition or concern. Add to this the fact that part-time, working, commuter students are extremely difficult to study, plus the fact that community colleges may often lack the institutional resources to support ongoing assessment and research efforts, and it becomes readily apparent why we know so little about… community colleges" (*The Increasing Importance of Community Colleges*, ¶3).

Subsequently, the estimates obtained from the 44 indicators are not generalizable to students who attend all types of PSIs in the U.S., and future research is necessary to understand the mental health experiences of students who attend different types of colleges, and community colleges in particular.

In addition to self-selection of institutions, students from each PSI self-selected to respond to the ACHA-NCHA online surveys. As mentioned in the discussion of the limitations of survey research, individuals who self-select to respond to surveys are known to differ from those who do not in a variety of ways. First, the literature overwhelmingly states they are more likely to be female (e.g., Eysenbach & Wyatt, 2002). A recent study of college students suggests that White and Asian American students may be more likely to self-select as well (Cranford et al., 2008). Furthermore, students who self-select to participate in survey research may differ from those who do not in terms of their health beliefs and health behaviors (Cranford et al.).

In summary, this study utilized a non-representative sample of U.S. institutions, and students from those institutions, though sampled using random or census methods, self-selected to complete the survey for which they were randomized. Because of the complex nature of this sample, the level to which students who completed the survey are representative of their particular institution remains unknown. As this was a pre-test of the revised ACHA-NCHA, and pre-tests are designed to gather preliminary information about the ways in which a survey performs in a sample and with methodology similar to that which will be employed with its official implementation, this is less of an issue than if these data were going to be used to make concrete estimates for the population of interest. As such, the issues of representativeness of both the institutions and students surveyed limits the generalizability of finding to U.S. college students, though some of these limitations are balanced by the large number of students who participated in this study. Ultimately, future research should sample a more comprehensive selection of U.S. PSIs, including 2-year schools, non-research institutions, private institutions, minority serving institutions, and more, and PSIs are encouraged to use appropriate incentives to increase response proportions to every extent possible.

Limitations of analyses and results

The final area of this study's limitations is that of the analyses used and subsequently the results reported. First, as is often common practice in these sorts of analyses (Garson, n.d.), the issue of the cluster sampling was ignored—but only after initial results showed no emergent patterns of student responses across institutions. However, future studies, and particularly those that utilize a more diverse sample of PSIs, must be cautioned against ignoring these issues without such preliminary analyses.

Additionally, if national prevalence estimates for these mental health conditions and experiences are to be generated and published as the "gold standard" for college student mental health, researchers will certainly want to address the issues of complex sampling.

Additionally, as has been suggested throughout this chapter, many of the results presented in Chapter Four were merely descriptive in nature and did not test any hypotheses. For example, no inferences can be made about differences observed by gender, race/ethnicity, and year in school for many of the results. Some may be made in terms of stress level, recency of negative affect, and diagnosis/treatment with depression in the last 12 months because they were entered into regression models when testing Hypothesis 4B. However, all other results are purely descriptive and readers must understand that though some estimates appear to differ between groups, these differences have been evaluated neither for their statistical nor their practical significance. Future studies are encouraged to more thoroughly examine these between-group differences.

As mentioned above, this research utilized ordinary least squares and logistic regression models to evaluate the extent to which the revised ACHA-NCHA mental health indicators demonstrated evidence of construct-related validity. Regression is one of the most commonly used modeling strategy in the social sciences (Field, 2005); yet, path analysis and structural equation modeling are emergent and relevant analytical techniques that may be used to more fully understand the direct and indirect relationships between and among variables. Thus, future research is suggested to utilize such analytical methods to more fully understand and evaluate these relationships.

Lastly, as was mentioned previously in this chapter in the discussion of results for Hypotheses 4A and 5, there are many limitations to the comparisons made between the

ACHA-NCHA estimates and those collected from the NCS-R. The instruments and populations are not directly comparable, and all results reported are more descriptive than inferential in nature. Presser and colleagues (2004) suggest that it is extremely important to validate estimates generated for phenomena across different versions of surveys to determine which are more accurate; however, they also highlight that validation data are rarely available and are themselves subject to error. This appears to be the case for this research study; though the NCS-R data are perhaps the best available for these types of comparisons, the extent to which true comparisons can be made are quite limited. Perhaps in the future, the ACHA-NCHA will be used in a nationally representative study and estimates collected for mental health indicators will become the gold standard against which many other smaller-scale surveys validate their results. In the meantime, as suggested by DeVellis (2003), "Although imperfect measurement may be better than no measurement at all in some situations, we should *recognize* when our measurement procedures are flawed and temper our conclusions accordingly" (p. 12).

Directions for Future Research

The results of this study suggest several potential directions for future research. First, qualitative techniques such as cognitive interviews (e.g., Conrad & Blair, 2004) should be used to elicit more comprehensive and detailed information about how students process each of the revised mental health indicators on the ACHA-NCHA. This qualitative information would add further evidence of validity and would provide a wealth of additional information to that collected from this study, particularly for those items measuring latent constructs. For example, ten of the items on the survey ask

students to identify if a variety of events have been "difficult to handle" in the last 12 months. Qualitative data would be useful in determining exactly what constitutes difficulty handling something (e.g., loss of sleep, excessive stress) in this population. As suggested by Ouimet and colleagues (2004), when triangulation is employed to validate a survey, researchers obtain an array of data points that can be used to cross reference areas needing improvement and/or provide evidence of further construct- and criterion-related validity.

As mentioned throughout this dissertation, this study examined the reliability and validity of only a small subset of items on the revised ACHA-NCHA. The revised survey consists of more than 300 variables measuring six broad content areas: (1) health, health education, and safety; (2) alcohol, tobacco, and other drugs; (3) sexual health; (4) mental and physical health; (5) impediments to academic performance; and (6) demographic characteristics. Future studies evaluating the psychometric properties of the other variables and sections of the survey are needed to ensure the entire revised ACHA-NCHA survey is collecting reliable and valid data from its target population. Furthermore, other sections of the survey were also modified from the original, and the extent to which student response patterns have changed across versions for all sections must be assessed. As was demonstrated in this study, changes to the way questions were asked produced both statistically and practically significant differences in student responses for the mental health indicators. Similar analyses must be conducted for the other sections and comparable items to determine the extent of differences across other indicators of students' health status.

In addition to examining changes in other items and indicators on the revised ACHA-NCHA, the impact of specific changes should be examined to the extent possible. As highlighted earlier in this chapter, several of the revised items differed from the original items in multiple ways including the question stem, response options, and recall time period. Subsequently, it is not possible to determine the specific impact of each of the modifications on response patterns. Future studies are needed to determine how each small change impacts student responses to ensure questions are as reliable and valid as possible.

As is cited in the literature, validation of survey data is a continuous, ongoing process. As such, continued validation studies are necessary for the mental health indicators. In particular, criterion-related validity could be assessed to a more extensive degree through the comparison of prevalence estimates obtained from ACHA-NCHA with medical records. Although it is likely not feasible to conduct a nation-wide study of this kind, the ACHA-NCHA mental health data could be compared to medical records for a select group of campuses, which would certainly provide more extensive information in regards to criterion-related validity.

Future research is also necessary to better understand the relationships among mental health constructs, and between mental health indicators and other variables, which could both provide further evidence of construct-related validity and be used to plan appropriate prevention and education programming. For example, the initial relationships observed between stress, negative affect, and diagnosis with mental health conditions could be further explored using complex statistical analyses such as path analysis or structural equation modeling. In addition, once indicators associated with other health

concepts (e.g., alcohol use, sexual behavior) have been examined for their reliability and validity as suggested above, the relationship between mental health and other risk behaviors can be examined and evaluated. Moreover, demographic relationships initially observed in this study can be further evaluated to determine which groups are at highest risk for mental illness and concern on college campuses. The exploration of these relationships will assist faculty, staff, and administrators on college campuses to most appropriately address these issues on their campuses.

As the need to understand college student mental health and general health status continues to grow, certainly future studies are necessary that sample a larger, more generalizable range of PSIs and students. The Core Alcohol and Drug Study (Core Institute, 2007) utilized a multistage sampling methodology and analysis plan that allows its results to be generalized to the entire U.S. college student population. Such a study using the revised ACHA-NCHA would be invaluable to the field of college health, particularly because the NCHRBS—which was conducted in 1995—was the last nationally generalizable comprehensive health study conducted with college students. This type of study would allow researcher to examine differences across a variety of individual as well as institutional factors (e.g., public/private, 2-year/4-year, Carnegie classification), and could provide extensive information to help college and public health professionals prioritize need and develop effective primary, secondary, and tertiary prevention strategies with the ultimate goal of improving mental health outcomes in this population.

Conclusion

As the mental health crisis among college students continues to rise, it is imperative that colleges and universities across the country have an adequate understanding of what is occurring on their campuses. This research demonstrates that the revised ACHA-NCHA is likely to be an appropriate, useful, and valuable tool to collect prevalence data, assess need, monitor trends, and evaluate health promotion efforts among today's college students. Only when the problems associated with student distress are understood may they be impacted through prevention and intervention. Hence, the creation of this revised survey instrument, which has demonstrated preliminary evidence of reliability and validity, has tremendous potential to aid student health professionals in their efforts to assess and then optimize the health of college students on their campuses.

The American College Health Association (ACHA) has been dedicated to the health needs of students at colleges and universities since the 1920s and is the principal leadership organization for the field of college health. ACHA provides services, communications, and advocacy that helps its members to advance the health of their campus communities (ACHA, 2007c).

Three types of membership are available within the ACHA. *Institutional membership* is available for entire institutions of higher education (i.e., colleges and universities and their departments that serve students); *individual membership* is available for individual health care providers such as physicians, physician assistants, administrators, nurses, nurse practitioners, mental health professionals, health educators, dietitians and nutritionists, pharmacists, and students; and finally, *sustaining membership* is available for corporations and other nonprofit organizations (e.g., pharmaceutical companies, software companies, insurance companies) interested in becoming more connected to the field of college health. Currently, ACHA has more than 900 institutional members and more than 2,400 individual members (ACHA, 2007c).

Member institutions represent the diversity of the higher education community – two and four year, public and private, large and small. Individual members represent all types of professionals, as well as students, who are dedicated to the advancement and protection of college student health. In addition, through an affiliation with several national accrediting organizations, ACHA provides continuing education and contact hours for physicians, physicians assistants, nurses, health educators, psychologists and pharmacists.

Original version of the online American College Health Association-
National College Health Assessment (ACHA-NCHA)

National College Health Assessment

The following questions ask about various aspects of your health. This survey is completely voluntary. You may choose not to participate or not to answer any specific questions. You may skip any question you are not comfortable answering. The survey is confidential. E-mail contact information is destroyed before data are compiled to protect anonymity. Composite data will then be shared with your campus for use in health promotion activities.

**NAVIGATE WITH THE MOUSE OR PRESS THE TAB KEY AFTER EACH ENTRY
DO NOT USE THE ENTER KEY**

The first 10 questions ask about health, health education, and safety

1) Considering your age, how would you describe your general health?

○ Excellent ○ Very good ○ Good ○ Fair ○ Poor ○ Don't know

2) On which of the following health topics have you received information from your college or university in the last school year? (Select all that apply)

☐ Tobacco use prevention
☐ Alcohol and other drug use prevention
☐ Sexual assault/relationship violence prevention
☐ Violence prevention
☐ Injury prevention and safety
☐ Suicide prevention

☐ Pregnancy prevention
☐ AIDS/ HIV
☐ Sexually transmitted disease (STD) prevention
☐ Dietary behaviors and nutrition
☐ Physical activity and fitness
☐ None of the above

3) Use the scale below to record the BELIEVABILITY of each source of health information (Please mark the best response for each question to the right)

	Believable	Neither Believable Nor Unbelievable	Unbelievable
Leaflets, pamphlets, flyers	○	○	○
Campus newspaper articles	○	○	○
Health center medical staff	○	○	○
Health educators	○	○	○
Friends	○	○	○

Original version of the online American College Health Association-
National College Health Assessment (ACHA-NCHA)

Resident assistants/advisors	○	○	○
Parents	○	○	○

4) Use the scale below to record the BELIEVABILITY of each source of health information (Please mark the best response for each question to the right)

	Believable	Neither Believable Nor Unbelievable	Unbelievable
Religious Center	○	○	○
Television	○	○	○
Magazines	○	○	○
Campus peer educators	○	○	○
Faculty/coursework	○	○	○
Internet/world wide web	○	○	○
Other	○	○	○

5) Do you usually get health-related information from any of the following sources?

	No	Yes
Leaflets, pamphlets, flyers	○	○
Campus newspaper articles	○	○
Health center medical staff	○	○
Health educators	○	○
Friends	○	○
Resident assistants/advisors	○	○
Parents	○	○

6) Do you usually get health-related information from any of the following sources?

	No	Yes
Religious Center	○	○
Television	○	○
Magazines	○	○
Campus peer educators	○	○
Faculty/coursework	○	○
Internet/world wide web	○	○
Other	○	○

7) Within the last school year, how often did you: (Please mark the appropriate column for each row)

	N/A didn't do this within the last school year	Never	Rarely	Sometimes	Most of the time	Always
Wear a seatbelt when you rode in a car?	○	○	○	○	○	○
Wear a helmet when you rode a bicycle?	○	○	○	○	○	○
Wear a helmet when you rode a motorcycle	○	○	○	○	○	○
Wear a helmet when you were inline skating?	○	○	○	○	○	○

8) Within the last school year, were you:

Original version of the online American College Health Association-
National College Health Assessment (ACHA-NCHA)

	No	Yes
In a physical fight?	○	○
Physically assaulted (do not include sexual assault)?	○	○

9) Within the last school year, have you experienced:

	No	Yes
Verbal threats for sex against your will?	○	○
Sexual touching against your will?	○	○
Attempted sexual penetration (vaginal, anal, oral intercourse) against your will?	○	○
Sexual penetration (vaginal, anal, oral intercourse) against your will?	○	○

10) Within the last school year, have you been in a relationship that was:

	No	Yes
Emotionally abusive?	○	○
Physically abusive?	○	○
Sexually abusive?	○	○

Go to next question

Original version of the online American College Health Association-
National College Health Assessment (ACHA-NCHA)

National College Health Assessment

The next 15 questions ask about alcohol, tobacco, and drugs.

11) Within the last 30 days, on how many days did you use: (Mark one for each row)

	Never used	Have used, but not in last 30 days	1-2 days	3-5 days	6-9 days	10-19 days	20-29 days	All 30 days
Cigarettes	○	○	○	○	○	○	○	○
Cigars	○	○	○	○	○	○	○	○
Smokeless tobacco	○	○	○	○	○	○	○	○
Alcohol (beer, wine, liquor)	○	○	○	○	○	○	○	○
Marijuana (pot, hash, hash oil)	○	○	○	○	○	○	○	○

12) Within the last 30 days, on how many days did you use: (Mark one for each row)

	Never used	Have used, but not in last 30 days	1-2 days	3-5 days	6-9 days	10-19 days	20-29 days	All 30 days
Cocaine (crack, rock, freebase)	○	○	○	○	○	○	○	○
Amphetamines (diet pills, speed, meth, crank)	○	○	○	○	○	○	○	○
Rohypnol (roofies), GHB, or Liquid X (intentional use)	○	○	○	○	○	○	○	○
MDMA (Ecstasy, XTC, E, X, Adam)	○	○	○	○	○	○	○	○
Other drugs	○	○	○	○	○	○	○	○

Original version of the online American College Health Association-
National College Health Assessment (ACHA-NCHA)

13) Within the last 30 days, how often do you think the typical student at your school used: State your best estimate. (Mark one for each row)

	Never used	One or more days	Used daily
Cigarettes	○	○	○
Cigars	○	○	○
Smokeless tobacco	○	○	○
Alcohol (beer, wine, liquor)	○	○	○
Marijuana (pot, hash, hash oil)	○	○	○
Cocaine (crack, rock, freebase)	○	○	○
Amphetamines (diet pills, speed, meth, crank)	○	○	○
Rohypnol (roofies), GHB, or Liquid X (intentional use)	○	○	○
MDMA (Ecstasy, XTC, E, X, Adam)	○	○	○
Other drugs	○	○	○

One drink of alcohol beverage is defined as a 12 oz. beer, a 4 oz. glass of wine, a shot of liquor, or a mixed drink.

14) Within the last 30 days, did you: (Mark one for each row)

	Not applicable Don't drive	Not applicable Don't drink	No	Yes
Drive after drinking any alcohol at all	○	○	○	○
Drive after having 5 or more drinks	○	○	○	○

15) The last time you "partied"/socialized, how many hours did you drink alcohol? State your best estimate.

_____ Hours

16) The last time you "partied"/socialized, how many alcoholic drinks did you have? State your best estimate.

_____ Drinks

17) In the last two weeks, on how many occasions did you drink the same or more alcohol as indicated in item #16? State your best estimate.

Occasions

18) How many alcoholic drinks do you think the typical student at your school had the last time he/she "partied"/socialized?

Drinks

19) Think back over the last two weeks. How many times, if any, have you had five or more alcoholic drinks at a sitting?

◯ None   ◯ 3 times   ◯ 6 times   ◯ 9 or more times
◯ 1 time   ◯ 4 times   ◯ 7 times
◯ 2 times   ◯ 5 times   ◯ 8 times

(Please mark the appropriate column for each row)

20) During the last school year, if you "partied"/socialized, how often did you...

	Not applicable Don't drink	Always	Usually	Sometimes	Rarely	Never
Alternate non-alcoholic with alcoholic beverages	◯	◯	◯	◯	◯	◯
Determine, in advance, not to exceed a set number of drinks	◯	◯	◯	◯	◯	◯
Choose not to drink alcohol	◯	◯	◯	◯	◯	◯
Use a designated driver	◯	◯	◯	◯	◯	◯
Eat before						

Original version of the online American College Health Association-
National College Health Assessment (ACHA-NCHA)

and/or
during
drinking   ○   ○   ○   ○   ○   ○

**21) During the last school year, if you "partied"/socialized, how often did
you...**

	Not applicable Don't drink	Always	Usually	Sometimes	Rarely	Never
Have a friend let you know when you've had enough	○	○	○	○	○	○
Keep track of how many drinks you were having	○	○	○	○	○	○
Pace your drinks to 1 or fewer per hour	○	○	○	○	○	○
Avoid drinking games	○	○	○	○	○	○
Drink an alcoholic look-alike (non-alcoholic beer, punch etc.)	○	○	○	○	○	○

(Please mark the appropriate column for each row)

**22) If you drink alcohol, within the last school year, have you experienced any
of the following as a consequence of your drinking?**

Original version of the online American College Health Association-
National College Health Assessment (ACHA-NCHA)

	Not applicable Don't drink	No	Yes
Physically injured yourself	○	○	○
Physically injured another person	○	○	○
Been involved in a fight	○	○	○
Did something you later regretted	○	○	○
Forgot where you were or what you did	○	○	○
Force/threat for sex	○	○	○
Had unprotected sex	○	○	○

23) Within the last 30 days, what percent of students at your school used Cigarettes? State your best estimate.

_____ Percent

24) Within the last 30 days, what percent of students at your school used Alcohol? State your best estimate.

_____ Percent

25) Within the last 30 days, what percent of students at your school used Rohypnol or GHB? State your best estimate.

_____ Percent

[Previous Page]　　[Go to next question]

National College Health Assessment

The next 11 questions ask about sex behavior, perceptions, and contraception.

26) Within the last school year, with how many partners, if any, have you had sex (oral, vaginal, or anal)?

| Number of Partners

27) Within the last school year, were your sexual partner(s), if any,

○ N/A   ○ Male   ○ Female   ○ Both Male and Female

28) Within the last school year, with how many partners do you think the typical student at your school has had sex (oral, vaginal, or anal)?

| Number of Partners

(Please mark the appropriate column for each row)

29) Within the last 30 days, if you are sexually active, how many times did you have:

	Never did this sexual activity	Have not done this during last 30 days	1-2 times	3-4 times	5-6 times	7-8 times	9-10 times	11 or more times
Oral sex?	○	○	○	○	○	○	○	○
Vaginal Intercourse?	○	○	○	○	○	○	○	○
Anal Intercourse?	○	○	○	○	○	○	○	○

(Please mark the appropriate column for each row)

30) How many times within the last 30 days do you think the typical student at your school has had:

	0 times	1-2 times	3-4 times	5-6 times	7-8 times	9-10 times	11 or more times
Oral sex?	○	○	○	○	○	○	○

Vaginal
Intercourse?   ○   ○   ○   ○   ○   ○   ○

Anal
Intercourse?   ○   ○   ○   ○   ○   ○   ○

(Please mark the appropriate column for each row)

Condom Use
**31) Within the last 30 days, if you are sexually active, how often did you or
your partner(s) use a condom during:**

	Never did this sexual activity	Have not done this during last 30 days	Never	Rarely	Sometimes	Mostly	Always
Oral sex?	○	○	○	○	○	○	○
Vaginal Intercourse?	○	○	○	○	○	○	○
Anal Intercourse?	○	○	○	○	○	○	○

(Please mark the appropriate column for each row)

Condom Use
**32) Within the last 30 days, how often do you think the typical student at your
school has used a condom during:**

	Typical student not participate	Never	Rarely	Sometimes	Mostly	Always
Oral sex?	○	○	○	○	○	○
Vaginal Intercourse?	○	○	○	○	○	○
Anal Intercourse?	○	○	○	○	○	○

(Please mark the appropriate column for each row)

33) If you are sexually active, did you use a condom the last time you had:

	Never did this sexual activity	No	Yes	Don't know Don't remember
Oral sex?	○	○	○	○
Vaginal Intercourse?	○	○	○	○
Anal Intercourse?	○	○	○	○

34) If you have had vaginal intercourse, what method did you or your partner use to prevent pregnancy the last time? (Select all that apply)

☐ Have not had vaginal intercourse
☐ Birth control pills
☐ Depo Provera (shots)
☐ Norplant (implant)
☐ Condoms (male or female)
☐ Diaphragm/Cervical cap/Sponge

☐ Spermicide (e.g. foam)
☐ Fertility awareness (calendar, mucous, basal body temperature)
☐ Withdrawal
☐ Other method
☐ Nothing

35) Within the last school year, if you are sexually active, have you or your partner(s) used emergency contraception ("morning after pill")?

○ No   ○ Yes   ○ Don't know   ○ Not sexually active

36) Within the last school year, have you unintentionally become pregnant or gotten someone else pregnant?

○ Have not had vaginal intercourse within the last school year
○ No
○ Yes
○ Don't know

The next 5 questions ask about weight, nutrition, and exercise.

37) How do you describe your weight?

○ Very underweight
○ Slightly underweight
○ About the right weight
○ Slightly overweight
○ Very overweight

38) Are you trying to do any of the following about your weight?

○ I am not trying to do anything about my weight
○ Stay the same weight

Original version of the online American College Health Association-
National College Health Assessment (ACHA-NCHA)

○ Lose weight
○ Gain weight

39) Within the last 30 days, did you do any of the following? (select all that apply)

☐ Exercise to lose weight
☐ Diet to lose weight
☐ Vomit or take laxatives to lose weight
☐ Take diet pills to lose weight
☐ I didn't do any of the above

40) How many servings of fruits and vegetables do you usually have per day (1 serving=1 medium piece of fruit, 1/2 cup chopped, cooked or canned fruit/vegetables, 3/4 cup fruit/vegetable juice, small bowl of salad greens, or 1/2 cup dried fruit)?

○ I don't eat fruits and vegetables
○ 1-2
○ 3-4
○ 5 or more

(Please mark the appropriate column for each row)

41) On how many of the past 7 days did you:

	0 days	1 day	2 days	3 days	4 days	5 days	6 days	7 days
Participate in vigorous exercise for at least 20 minutes or moderate exercise for at least 30 minutes?	○	○	○	○	○	○	○	○
Do exercises to strength or tone your muscles, such as push-ups, sit-ups, or weight lifting?	○	○	○	○	○	○	○	○
Get enough sleep so that you felt rested when you woke up in the morning?	○	○	○	○	○	○	○	○

The next 10 questions ask about mental and physical health.

(Please mark the appropriate column for each row)

42) Within the last school year how many times have you:

Original version of the online American College Health Association-
National College Health Assessment (ACHA-NCHA)

	Never	1-2 times	3-4 times	5-6 times	7-8 times	9-10 times	11 or more times
Felt things were hopeless	○	○	○	○	○	○	○
Felt overwhelmed by all you had to do	○	○	○	○	○	○	○
Felt exhausted (not from physical activity)	○	○	○	○	○	○	○
Felt very sad	○	○	○	○	○	○	○
Felt so depressed that it was difficult to function	○	○	○	○	○	○	○
Seriously considered attempting suicide	○	○	○	○	○	○	○
Attempted suicide	○	○	○	○	○	○	○

Please press the *Go to next question* button after answering the following question.

43) Have you ever been diagnosed with depression?

○ Yes   ○ No

[Previous Page]   [Go to next question]

Original version of the online American College Health Association-
National College Health Assessment (ACHA-NCHA)

National College Health Assessment

44) If Yes:

	No	Yes
Have you been diagnosed with depression within the last school year?	○	○
Are you currently in therapy for depression?	○	○
Are you currently taking medication for depression?	○	○

45) Have you:

	No	Yes	Don't Know
Been vaccinated against hepatitis B?	○	○	○
Been vaccinated against meningococcal disease (meningococcal meningitis)?	○	○	○
Been vaccinated against varicella (chicken pox)?	○	○	○
Been vaccinated with measles, mumps, rubella (2 shots)?	○	○	○
Been vaccinated against influenza (the flu) in the last year?	○	○	○
Had a dental exam and cleaning in the last year?	○	○	○
(Males) Performed testicular self exam in the last month?	○	○	○
(Females) Performed breast self exam in the last month?	○	○	○
(Females) Had a routine gynecological exam in the last year?	○	○	○
Had your blood pressure checked in the last 2 years?	○	○	○
Had your cholesterol checked in the last 5 years?	○	○	○
Used sunscreen daily?	○	○	○

46) Have you ever been tested for HIV infection?

○ No   ○ Yes   ○ Don't know

47) Within the last school year, have you had any of the following?

	No	Yes
Allergy problems	○	○

Original version of the online American College Health Association-
National College Health Assessment (ACHA-NCHA)

	No	Yes
Anorexia	○	○
Anxiety Disorder	○	○
Asthma	○	○
Bulimia	○	○
Chronic Fatigue Syndrome	○	○
Depression	○	○
Diabetes	○	○
Endometriosis	○	○
Genital Herpes	○	○
Genital warts/HPV	○	○
Hepatitis B or C	○	○
High blood pressure	○	○
High cholesterol	○	○
HIV infection	○	○

48) Within the last school year, have you had any of the following?

	No	Yes
Repetitive stress injury (e.g. carpal tunnel syndrome)	○	○
Seasonal Affect Disorder	○	○
Substance abuse problem	○	○
Back pain	○	○
Broken bone/fracture	○	○
Bronchitis	○	○
Chlamydia	○	○
Ear Infection	○	○
Gonorrhea	○	○
Mononucleosis	○	○
Pelvic Inflammatory Disease	○	○
Sinus Infection	○	○

Appendix B1:
Original version of the online American College Health Association-
National College Health Assessment (ACHA-NCHA)

Strep throat	○	○
Tuberculosis	○	○

49) Have you ever been diagnosed with any of the following?

	No	Yes
Allergy problems	○	○
Anorexia	○	○
Anxiety Disorder	○	○
Asthma	○	○
Bulimia	○	○
Chronic Fatigue Syndrome	○	○
Depression	○	○
Diabetes	○	○
Endometriosis	○	○
Genital Herpes	○	○
Genital warts/HPV	○	○
Hepatitis B or C	○	○
High blood pressure	○	○
High cholesterol	○	○
HIV infection	○	○

50) Have you ever been diagnosed with any of the following?

	No	Yes
Repetitive stress injury (e.g. carpal tunnel syndrome)	○	○
Seasonal Affect Disorder	○	○
Substance abuse problem	○	○
Back pain	○	○
Broken bone/fracture	○	○
Bronchitis	○	○

Original version of the online American College Health Association-
National College Health Assessment (ACHA-NCHA)

Chlamydia	○	○
Ear Infection	○	○
Gonorrhea	○	○
Mononucleosis	○	○
Pelvic Inflammatory Disease	○	○
Sinus Infection	○	○
Strep throat	○	○
Tuberculosis	○	○

The next 4 questions ask about impediments to academic performance.

51) Within the last school year, have any of the following affected your academic performance? (Please select the most serious outcome for each item below)

	This did not happen to me/not applicable	I have experienced this issue but my academics have not been affected	Received a lower grade on an exam or important project	Received a lower grade in the course	Received an incomplete or dropped the course
Alcohol use	○	○	○	○	○
Allergies	○	○	○	○	○
Assault (physical)	○	○	○	○	○
Assault (sexual)	○	○	○	○	○
Attention Deficit Disorder	○	○	○	○	○

52) Within the last school year, have any of the following affected your academic performance? (Please select the most serious outcome for each item below)

	This did not happen to me/not applicable	I have experienced this issue but my academics have not been affected	Received a lower grade on an exam or important project	Received a lower grade in the course	Received an incomplete or dropped the course
Cold/Flu/Sore	○	○	○	○	○

Original version of the online American College Health Association-
National College Health Assessment (ACHA-NCHA)

throat

Concern for a troubled friend or family member	○	○	○	○	○
Chronic illness (diabetes, asthma, etc.)	○	○	○	○	○
Chronic pain	○	○	○	○	○
Death of a friend or family member	○	○	○	○	○

53) Within the last school year, have any of the following affected your academic performance? (Please select the most serious outcome for each item below)

	This did not happen to me/not applicable	I have experienced this issue but my academics have not been affected	Received a lower grade on an exam or important project	Received a lower grade in the course	Received an incomplete or dropped the course
Depression/Anxiety Disorder/Seasonal Affective Disorder	○	○	○	○	○
Drug use	○	○	○	○	○
Eating disorder/problem	○	○	○	○	○
HIV Infection	○	○	○	○	○
Injury	○	○	○	○	○
Internet use/computer games	○	○	○	○	○
Learning disability	○	○	○	○	○

54) Within the last school year, have any of the following affected your academic performance? (Please select the most serious outcome for each item below)

	This did not	I have experienced this issue but my	Received a lower grade on	Received a lower	Received an incomplete

Original version of the online American College Health Association-
National College Health Assessment (ACHA-NCHA)

	happen to me/not applicable	academics have not been affected	an exam or important project	grade in the course	or dropped the course
Mononucleosis	○	○	○	○	○
Pregnancy (yours or your partner's)	○	○	○	○	○
Relationship difficulty	○	○	○	○	○
Sexually transmitted disease	○	○	○	○	○
Sinus infection/ear infection/bronchitis/strep throat	○	○	○	○	○
Sleep difficulty	○	○	○	○	○
Stress	○	○	○	○	○
Other	○	○	○	○	○

The last questions ask about demographic characteristics.

55) How old are you?

_____ Years

56) What is your sex?

○ Female
○ Male

The next two questions ask about your height. For example if your height is 5 foot, 7 inches, please indicate "5" in question 57 and "7" in question 58.

57) What is your height in feet?

_____

58) and inches?

59) What is your weight in pounds?

_____ Pounds

60) Year in school:

Original version of the online American College Health Association-
National College Health Assessment (ACHA-NCHA)

Select One ▢

61) Are you a full-time student?

○ Yes ○ No

62) How do you usually describe yourself? (Mark all that apply)

☐ White-not Hispanic (includes Middle Eastern)
☐ Black-not Hispanic
☐ Hispanic or Latino
☐ Asian or Pacific Islander
☐ American Indian or Alaskan Native
☐ Other

63) Are you an international student?

○ Yes ○ No

64) What is your current relationship status?

○ Single
○ Married/domestic partner
○ Engaged or committed dating relationship
○ Separated
○ Divorced
○ Widowed

65) Which of the following best describes you?

○ Heterosexual ○ Gay/Lesbian ○ Bisexual ○ Transgendered ○ Unsure

66) Where do you currently live?

○ Campus residence hall
○ Fraternity or sorority house
○ Other university/college housing
○ Off-campus housing
○ Parent/guardian's home
○ Other

67) Are you a member of a social fraternity or sorority? (National Interfraternity Conference, National Panhellenic Conference, or National Pan-Hellenic Council)

○ Yes ○ No

68) How many hours a week do you work for pay?

Select One ▢

Original version of the online American College Health Association-
National College Health Assessment (ACHA-NCHA)

69) If you have a credit card(s) how much total credit card debt did you carry last month? That is, what was the total unpaid balance on all of your cards (that you are responsible for paying)?

Select One

70) What is your approximate cumulative grade average?

○ A   ○ B   ○ C   ○ D/F   ○ N/A

71) How many hours a week do you volunteer?

Select One

72) Do you have any kind of health insurance (including prepaid plans such as HMOs-health maintenance organizations)?

○ Yes   ○ No   ○ Not sure

Thank you for taking the time and thought to complete this survey. We appreciate your participation!

[Previous Page]   [Submit Survey]

Online Survey Software powered by Vovici.

Revised version of the online American College Health Association-
National College Health Assessment (ACHA-NCHA)

National College Health Assessment

The following questions ask about various aspects of your health. This survey is completely voluntary. You may choose not to participate or not to answer any specific questions. You may skip any question you are not comfortable answering. The survey is confidential. E-mail contact information is destroyed before data are compiled to protect anonymity. Composite data will then be shared with your campus for use in health promotion activities.

**NAVIGATE WITH THE MOUSE OR PRESS THE TAB KEY AFTER EACH ENTRY
DO NOT USE THE ENTER KEY**

Health, Health Education, and Safety

1) How would you describe your general health?

       ○ Excellent   ○ Very good   ○ Good   ○ Fair   ○ Poor   ○ Don't know

2) Have you received information on the following topics from your college or university? (Please mark the appropriate column for each row)

	No	Yes
Alcohol and other drug use	○	○
Cold/Flu/Sore throat	○	○
Depression/Anxiety	○	○
Eating disorders	○	○
Grief and loss	○	○
How to help others in distress	○	○
Injury and violence prevention	○	○
Nutrition	○	○
Physical activity	○	○
Pregnancy prevention	○	○
Problem use of Internet/computer games	○	○
Relationship difficulties	○	○
Sexual assault/Relationship violence prevention	○	○
Sexually transmitted disease/infection (STD/I) prevention	○	○
Sleep difficulties	○	○

Revised version of the online American College Health Association-
National College Health Assessment (ACHA-NCHA)

	No	Yes
Stress reduction	○	○
Suicide prevention	○	○
Tobacco use	○	○

3) Are you interested in receiving information on the following topics from your college or university? (Please mark the appropriate column for each row)

	No	Yes
Alcohol and other drug use	○	○
Cold/Flu/Sore throat	○	○
Depression/Anxiety	○	○
Eating disorders	○	○
Grief and loss	○	○
How to help others in distress	○	○
Injury and violence prevention	○	○
Nutrition	○	○
Physical activity	○	○
Pregnancy prevention	○	○
Problem use of Internet/computer games	○	○
Relationship difficulties	○	○
Sexual assault/Relationship violence prevention	○	○
Sexually transmitted disease/infection (STD/I) prevention	○	○
Sleep difficulties	○	○
Stress reduction	○	○
Suicide prevention	○	○
Tobacco use	○	○

4) Within the last 12 months, how often did you: (Please mark the appropriate column for each row)

N/A, did
not do

Revised version of the online American College Health Association-
National College Health Assessment (ACHA-NCHA)

	this activity within the last 12 months	Never	Rarely	Sometimes	Most of the time	Always
Wear a seatbelt when you rode in a car?	○	○	○	○	○	○
Wear a helmet when you rode a bicycle?	○	○	○	○	○	○
Wear a helmet when you rode a motorcycle?	○	○	○	○	○	○
Wear a helmet when you were inline skating?	○	○	○	○	○	○

5) Within the last 12 months: (Please mark the appropriate column for each row)

	No	Yes
Were you in a physical fight?	○	○
Were you physically assaulted (do not include sexual assault)?	○	○
Were you verbally threatened?	○	○
Were you sexually touched without your consent?	○	○
Was sexual penetration attempted (vaginal, anal, oral) without your consent?	○	○
Were you sexually penetrated (vaginal, anal, oral) without your consent?	○	○
Were you a victim of stalking (e.g., waiting for you outside your classroom, residence hall, or office; repeated emails/phone calls)?	○	○

6) Within the last 12 months, have you been in an intimate (coupled/partnered) relationship that was: (Please mark the appropriate

Revised version of the online American College Health Association-
National College Health Assessment (ACHA-NCHA)

column for each row)

	No	Yes
Emotionally abusive? (e.g., called derogatory names, yelled at, ridiculed)	○	○
Physically abusive? (e.g., kicked, slapped, punched)	○	○
Sexually abusive? (e.g., forced to have sex when you didn't want it, forced to perform or have an unwanted sexual act performed on you)	○	○

7) Do you feel safe: (Please mark the appropriate column for each row)

	No	Yes
On campus?	○	○
Off Campus?	○	○

Go to next question

Revised version of the online American College Health Association-
National College Health Assessment (ACHA-NCHA)

National College Health Assessment

Alcohol, Tobacco, and Drugs

8) Within the last 30 days, on how many days did you use: (Please mark the appropriate column for each row)

	Never used	Have used, but not in last 30 days	1-2 days	3-5 days	6-9 days	10-19 days	20-29 days	Used daily
Cigarettes	○	○	○	○	○	○	○	○
Cigars, little cigars, bidis, hookah	○	○	○	○	○	○	○	○
Smokeless tobacco	○	○	○	○	○	○	○	○
Alcohol (beer, wine, liquor)	○	○	○	○	○	○	○	○
Marijuana (pot, weed)	○	○	○	○	○	○	○	○
Cocaine (crack, rock, freebase)	○	○	○	○	○	○	○	○
Methamphetamine (crystal meth, ice, crank)	○	○	○	○	○	○	○	○
Other amphetamines (diet pills, bennies)	○	○	○	○	○	○	○	○
Sedatives (downers, ludes)	○	○	○	○	○	○	○	○
Hallucinogens (LSD, PCP)	○	○	○	○	○	○	○	○
Steroids	○	○	○	○	○	○	○	○
Opiates (heroin, smack)	○	○	○	○	○	○	○	○
Inhalants (glue, solvents, gas)	○	○	○	○	○	○	○	○
MDMA (Ecstacy)	○	○	○	○	○	○	○	○

Revised version of the online American College Health Association-
National College Health Assessment (ACHA-NCHA)

Other club drugs (GHB, Ketamine, Rohypnol)	○	○	○	○	○	○	○	○
Other illegal drugs	○	○	○	○	○	○	○	○

9) Within the last 30 days, how often do you think the typical student at your school used: (State your best estimate; Please mark the appropriate column for each row)

	Never used	Have used, but not in last 30 days	1-2 days	3-5 days	6-9 days	10-19 days	20-29 days	Used daily
Cigarettes	○	○	○	○	○	○	○	○
Cigars, little cigars, bidis, hookah	○	○	○	○	○	○	○	○
Smokeless tobacco	○	○	○	○	○	○	○	○
Alcohol (beer, wine, liquor)	○	○	○	○	○	○	○	○
Marijuana (pot, weed)	○	○	○	○	○	○	○	○
Cocaine (crack, rock, freebase)	○	○	○	○	○	○	○	○
Methamphetamine (crystal meth, ice, crank)	○	○	○	○	○	○	○	○
Other amphetamines (diet pills, bennies)	○	○	○	○	○	○	○	○
Sedatives (downers, ludes)	○	○	○	○	○	○	○	○
Hallucinogens (LSD, PCP)	○	○	○	○	○	○	○	○
Steroids	○	○	○	○	○	○	○	○
Opiates (heroin, smack)	○	○	○	○	○	○	○	○
Inhalants (glue, solvents, gas)	○	○	○	○	○	○	○	○

MDMA (Ecstacy)	○	○	○	○	○	○	○	○
Other club drugs (GHB, Ketamine, Rohypnol)	○	○	○	○	○	○	○	○
Other illegal drugs	○	○	○	○	○	○	○	○

One drink of alcohol is defined as a 12 oz. can or bottle of beer or wine cooler, a 4 oz. glass of wine, or a shot of liquor straight or in a mixed drink.

10) The last time you "partied"/socialized how many drinks of alcohol did you have? (If you did not drink, please write 0)

Drinks

11) The last time you "partied"/socialized, over how many hours did you drink? (If you did not drink, please write 0)

Hours

12) Did you drive after the last time you "partied"/socialized?

○ No   ○ Yes

13) How many drinks of alcohol do you think the typical student at your school had the last time he/she "partied"/socialized?

_____ Drinks

14) Over the last two weeks, how many times have you had five or more drinks of alcohol at a sitting?

○ N/A, Don't drink
○ None
○ 1 time
○ 2 times
○ 3 times
○ 4 times
○ 5 times
○ 6 times
○ 7 times
○ 8 times
○ 9 times
○ 10 or more times

Revised version of the online American College Health Association-
National College Health Assessment (ACHA-NCHA)

15) During the last 12 months, when you "partied"/socialized, how often did you: (Please mark the appropriate column for each row)

	N/A, Don't drink	Never	Rarely	Sometimes	Most of the time	Always
Alternate non-alcoholic with alcoholic beverages	○	○	○	○	○	○
Avoid drinking games	○	○	○	○	○	○
Choose not to drink alcohol	○	○	○	○	○	○
Determine, in advance, not to exceed a set number of drinks	○	○	○	○	○	○
Eat before and/or during drinking	○	○	○	○	○	○
Have a friend let you know when you have had enough	○	○	○	○	○	○
Keep track of how many drinks you were having	○	○	○	○	○	○
Pace your drinks to 1 or fewer per hour	○	○	○	○	○	○
Stay with the same group of friends the entire time you were drinking	○	○	○	○	○	○
Stick with only one kind of alcohol when drinking	○	○	○	○	○	○
Use a designated driver	○	○	○	○	○	○

16) Within the last 12 months, how many times have you experienced any of the following as a consequence of your drinking? (Please mark the

appropriate column for each row)

	N/A, Don't drink	0 times	1 time	2 times	3 times	4 or more times
Did something you later regretted	○	○	○	○	○	○
Forgot where you were or what you did	○	○	○	○	○	○
Got in trouble with the police	○	○	○	○	○	○
Had sex with someone without giving your consent	○	○	○	○	○	○
Had sex with someone without getting their consent	○	○	○	○	○	○
Had unprotected sex	○	○	○	○	○	○
Physically injured yourself	○	○	○	○	○	○
Physically injured another person	○	○	○	○	○	○
Seriously considered suicide	○	○	○	○	○	○

17) In the last 12 months, have you taken any of the following prescriptions drugs that were not prescribed to you? (Please mark the appropriate column for each row)

	No	Yes
Antidepressants (e.g., Celexa, Lexapro, Prozac, Wellbutrin, Zoloft)	○	○
Erectile dysfunction drugs (e.g., Viagra, Cialis, Levitra)	○	○
Pain killers (e.g., OxyContin, Vicodin, Codeine)	○	○
Sedatives (e.g., Xanax, Valium)	○	○
Stimulants (e.g., Ritalin, Adderall)	○	○

[Previous Page]   [Go to next question]

Revised version of the online American College Health Association-
National College Health Assessment (ACHA-NCHA)

National College Health Assessment

Sex Behavior and Contraception

18) Within the last 12 months, with how many partners have you had oral sex, vaginal intercourse, or anal intercourse?

Number of Partners

19) Within last 12 months, did you have sexual partner(s) who were: (Please mark the appropriate column for each row)

	No	Yes
Female	○	○
Male	○	○
Transgender	○	○

20) Within the last 30 days, did you have: (Please mark the appropriate column for each row)

	No, have never done this sexual activity	No, have done this sexual activity in the past but not in the last 30 days	Yes
Oral sex?	○	○	○
Vaginal intercourse?	○	○	○
Anal intercourse?	○	○	○

21) Within the last 30 days, how often did you or your partner(s) use a condom or other protective barrier (e.g., male condom, female condom, dam, glove) during: (Please mark the appropriate column for each row)

	N/A, Never did this sexual activity	Have not done this sexual activity during the last 30 days	Never	Rarely	Sometimes	Most of the time	Always
Oral sex?	○	○	○	○	○	○	○
Vaginal							

intercourse? ○   ○   ○   ○   ○   ○   ○

Anal
intercourse? ○   ○   ○   ○   ○   ○   ○

22) What method of birth control did you or your partner use to prevent pregnancy the last time you had vaginal intercourse?

- ○ N/A, Have not had vaginal intercourse that could result in a pregnancy
- ○ N/A, Did not want to prevent pregnancy
- ○ Sterilization (e.g., hysterectomy, tubes tied, or vasectomy)
- ○ Birth control pills (monthly or extended cycle)
- ○ Birth control shots
- ○ Birth control implants
- ○ Birth control patch
- ○ Cervical ring
- ○ Intrauterine device (IUD)
- ○ Male condom only
- ○ Female condom only
- ○ Condom and another method
- ○ Diaphragm or cervical cap
- ○ Contraceptive sponge
- ○ Spermicide (e.g., foam, jelly, cream) only
- ○ Fertility awareness (e.g., calendar, mucous, basal body temperature)
- ○ Withdrawal
- ○ Other method
- ○ Nothing
- ○ Don't know/Not sure

23) Within the last 12 months, have you or your partner(s) used emergency contraception ("morning after pill")?

- ○ N/A, Have not had vaginal intercourse in the last 12 months
- ○ No
- ○ Yes
- ○ Don't know

24) Within the last 12 months, have you or your partner(s) become pregnant?

- ○ N/A, Have not had vaginal intercourse in the last 12 months
- ○ No
- ○ Yes, unintentionally
- ○ Yes, intentionally
- ○ Don't know

Weight, Nutrition, and Excercise

25) How do you describe your weight?

- ○ Very underweight
- ○ Slightly underweight
- ○ About the right weight
- ○ Slightly overweight
- ○ Very overweight

26) Are you trying to do any of the following about your weight?

- ○ I am not trying to do anything about my weight
- ○ Stay the same weight
- ○ Lose weight
- ○ Gain weight

27) How many servings of fruits and vegetables do you usually have per day? (1 serving = 1 medium piece of fruit, ½ cup fresh, frozen, or canned fruits/vegetables, ¾ cup fruit/vegetable juice, 1 cup salad greens, or ¼ cup dried fruit)

- ○ 0 servings per day
- ○ 1-2 servings per day
- ○ 3-4 servings per day
- ○ 5 or more servings per day

28) On how many of the past 7 days did you: (Please mark the appropriate column for each row)

	0 days	1 day	2 days	3 days	4 days	5 days	6 days	7 days
Participate in moderate exercise (caused light sweating or moderate increase in your heart rate) for at least 30 minutes?	○	○	○	○	○	○	○	○
Participate in vigorous exercise (caused heavy sweating or large increases in breathing or heart rate) for at least 20 minutes?	○	○	○	○	○	○	○	○
Do exercises to strengthen or tone your muscles, such as push-ups, sit-ups, or weight lifting?	○	○	○	○	○	○	○	○

[Previous Page]   [Go to next question]

National College Health Assessment

Mental and Physical Health

29) Have you ever: (Please mark the appropriate column for each row)

	Never	Not in the last 12 months	In the last 12 months	In the last 30 days	In the last 2 weeks
Felt things were hopeless	○	○	○	○	○
Felt overwhelmed by all you had to do	○	○	○	○	○
Felt exhausted (not from physical activity)	○	○	○	○	○
Felt very lonely	○	○	○	○	○
Felt very sad	○	○	○	○	○
Felt so depressed that it was difficult to function	○	○	○	○	○
Felt overwhelming anxiety	○	○	○	○	○
Felt overwhelming anger	○	○	○	○	○
Intentionally cut, burned, bruised, or otherwise injured yourself	○	○	○	○	○
Seriously considered suicide	○	○	○	○	○
Attempted suicide	○	○	○	○	○

30) Within the last 12 months, have you been diagnosed or treated by a professional for any of the following? (Please mark the appropriate column for each row)

	No	Yes, diagnosed but not treated	Yes, treated with medication	Yes, treated with psychotherapy	Yes, treated with medication and psychotherapy	Yes, other treatment
Anorexia	○	○	○	○	○	○
Anxiety	○	○	○	○	○	○
Attention Deficit and Hyperactivity Disorder (ADHD)	○	○	○	○	○	○
Bipolar Disorder	○	○	○	○	○	○

Bulimia	○	○	○	○	○	○
Depression	○	○	○	○	○	○
Insomnia	○	○	○	○	○	○
Other sleep disorder	○	○	○	○	○	○
Obsessive Compulsive Disorder (OCD)	○	○	○	○	○	○
Panic attacks	○	○	○	○	○	○
Phobia	○	○	○	○	○	○
Schizophrenia	○	○	○	○	○	○
Substance abuse or addiction (alcohol or other drugs)	○	○	○	○	○	○
Other addiction (e.g., gambling, internet, sexual)	○	○	○	○	○	○
Other mental health condition	○	○	○	○	○	○

31) Have you ever been diagnosed (by a professional) with depression?

○ No   ○ Yes

32) Within the last 12 months, have any of the following been difficult to handle for you? (Please mark the appropriate column for each row)

	No	Yes
Academics	○	○
Career related issue	○	○
Death of a family member or friend	○	○
Family problems	○	○
Other relationships	○	○
Finances	○	○
Health problem of a family member or partner	○	○

Revised version of the online American College Health Association-
National College Health Assessment (ACHA-NCHA)

	No	Yes
Personal appearance	○	○
Personal health issue	○	○
Other	○	○

33) Have you ever received psychological or mental health services from any of the following? (Please mark the appropriate column for each row)

	No	Yes
Counselor/Therapist/Psychologist	○	○
Psychiatrist	○	○
Other medical provider (e.g., physician, nurse practitioner)	○	○
Minister/Priest/Rabbi/Other clergy	○	○

34) Have you ever received psychological or mental health services from your current college/university's Counseling or Health Service?

○ No   ○ Yes

35) If in the future you were having a personal problem that was really bothering you, would you consider seeking help from a mental health professional?

○ No   ○ Yes

36) Within the last 12 months, how would you rate the overall level of stress you have experienced?

○ No stress
○ Less than average stress
○ Average stress
○ More than average stress
○ Tremendous stress

37) Within the last 30 days, did you do any of the following? (Please mark the appropriate column for each row)

	No	Yes
Exercise to lose weight	○	○
Diet to lose weight	○	○
Vomit or take laxatives to lose weight	○	○
Take diet pills to lose weight	○	○

38) Have you: (Please mark the appropriate column for each row)

	No	Yes	Don't Know

	No	Yes	Don't Know
Had a dental exam and cleaning in the last 12 months?	○	○	○
(Males) Performed a testicular self exam in the last 30 days?	○	○	○
(Females) Performed a breast self exam in the last 30 days?	○	○	○
(Females) Had a routine gynecological exam in the last 12 months?	○	○	○
Used sunscreen regularly with sun exposure?	○	○	○
Ever been tested for Human Immunodeficiency Virus (HIV) infection?	○	○	○

39) Have you received the following vaccinations (shots)? (Please mark the appropriate column for each row)

	No	Yes	Don't Know
Hepatitis B	○	○	○
Human Papilloma Virus (HPV)	○	○	○
Influenza (the flu) in the last 12 months (shot or nasal mist)	○	○	○
Measles, Mumps, Rubella (2 shots)	○	○	○
Meningococcal disease (meningococcal meningitis)	○	○	○
Varicella (chicken pox)	○	○	○

40A) Within the last 12 months, have you been diagnosed or treated for any of the following? (Please mark the appropriate column for each row)

	No	Yes
Allergies	○	○
Asthma	○	○
Back pain	○	○
Broken bone/Fracture/Sprain	○	○
Bronchitis	○	○
Chlamydia	○	○
Diabetes	○	○
Ear infection	○	○
Endometriosis	○	○
Genital herpes	○	○
Genital warts/Human Papilloma Virus (HPV)	○	○
Gonorrhea	○	○
Hepatitis	○	○

Revised version of the online American College Health Association-
National College Health Assessment (ACHA-NCHA)

40B) Within the last 12 months, have you been diagnosed or treated for any of the following? (Please mark the appropriate column for each row)

	No	Yes
High blood pressure	○	○
High cholesterol	○	○
Human Immunodeficiency Virus (HIV)	○	○
Irritable Bowel Syndrome	○	○
Migraine headache	○	○
Mononucleosis	○	○
Pelvic Inflammatory Disease (PID)	○	○
Repetitive stress injury (e.g., carpal tunnel syndrome)	○	○
Sinus infection	○	○
Strep throat	○	○
Tuberculosis	○	○
Urinary tract infection	○	○

41) On how many of the past 7 days did you get enough sleep so that you felt rested when you woke up in the morning?

○ 0 days   ○ 1 day   ○ 2 days   ○ 3 days   ○ 4 days   ○ 5 days   ○ 6 days   ○ 7 days

42) People sometimes feel sleepy during the daytime. In the past 7 days, how much of a problem have you had with sleepiness (feeling sleepy, struggling to stay awake) during your daytime activities?

○ Not a problem at all
○ A little problem
○ More than a little problem
○ A big problem
○ A very big problem

43) In the past 7 days, how often have you: (Please mark the appropriate column for each row)

	0 days	1 day	2 days	3 days	4 days	5 days	6 days	7 days
Awakened too early in the morning and couldn't get back to sleep?	○	○	○	○	○	○	○	○
Felt tired, dragged out, or sleepy during the day?	○	○	○	○	○	○	○	○
Gone to bed because you just could not stay awake any longer?	○	○	○	○	○	○	○	○
Had an extremely hard time falling asleep?	○	○	○	○	○	○	○	○

[Previous Page]   [Go to next question]

Revised version of the online American College Health Association-
National College Health Assessment (ACHA-NCHA)

National College Health Assessment

Impediments to Academic Performance

44A) Within the last 12 months, have any of the following affected your academic performance? (Please select the most serious outcome for each item below)

	N/A, This did not happen to me	I have experienced this issue but my academics have not been <u>negatively</u> affected	Received a lower grade on an exam or important project	Received a lower grade in the course	Received an incomplete or dropped the course
Alcohol use	○	○	○	○	○
Allergies	○	○	○	○	○
Anxiety	○	○	○	○	○
Assault (physical)	○	○	○	○	○
Assualt (sexual)	○	○	○	○	○
Attention Deficit and Hyperactivity Disorder (ADHD)	○	○	○	○	○
Cold/Flu/Sore throat	○	○	○	○	○
Concern for a troubled friend or family member	○	○	○	○	○

44B) Within the last 12 months, have any of the following affected your academic performance? (Please select the most serious outcome for each item below)

	N/A, This did not happen to me	I have experienced this issue but my academics have not been <u>negatively</u> affected	Received a lower grade on an exam or important project	Received a lower grade in the course	Received an incomplete or dropped the course

Revised version of the online American College Health Association-
National College Health Assessment (ACHA-NCHA)

Chronic health problem or serious illness (e.g., diabetes, asthma, cancer)	○	○	○	○	○
Chronic pain	○	○	○	○	○
Death of a friend or family member	○	○	○	○	○
Depression	○	○	○	○	○
Discrimination (e.g., homophobia, racism, sexism)	○	○	○	○	○
Drug use	○	○	○	○	○
Eating disorder/problem	○	○	○	○	○
Finances	○	○	○	○	○

44C) Within the last 12 months, have any of the following affected your academic performance? (Please select the most serious outcome for each item below)

	N/A, This did not happen to me	I have experienced this issue but my academics have not been <u>negatively</u> affected	Received a lower grade on an exam or important project	Received a lower grade in the course	Received an incomplete or dropped the course
Gambling	○	○	○	○	○
Homesickness	○	○	○	○	○
Injury (fracture, sprain, strain, cut)	○	○	○	○	○
Internet use/computer games	○	○	○	○	○
Learning disability	○	○	○	○	○
Participation in extracurricular activities (e.g.,					

campus clubs, organizations, athletics)	○	○	○	○	○
Pregnancy (yours or your partner's)	○	○	○	○	○
Relationship difficulties	○	○	○	○	○

44D) Within the last 12 months, have any of the following affected your academic performance? (Please select the most serious outcome for each item below)

	N/A, This did not happen to me	I have experienced this issue but my academics have not been negatively affected	Received a lower grade on an exam or important project	Received a lower grade in the course	Received an incomplete or dropped the course
Roommate difficulties	○	○	○	○	○
Sexually transmitted disease/infection (STD/I)	○	○	○	○	○
Sinus infection/Ear infection/Bronchitis/Strep throat	○	○	○	○	○
Sleep difficulties	○	○	○	○	○
Stress	○	○	○	○	○
Work	○	○	○	○	○
Other (please specify in "Additional Comments" box below)	○	○	○	○	○

Additional comments:

_____

Demographic Characteristics

45) How old are you?

_____
        Years

Revised version of the online American College Health Association-
National College Health Assessment (ACHA-NCHA)

46) What is your gender?

○ Female
○ Male
○ Transgender

47) What is your sexual orientation?

○ Heterosexual   ○ Gay/Lesbian   ○ Bisexual   ○ Unsure

The next two questions ask about your height. For example if your height is 5 foot, 7 inches, please indicate "5" in question 48A and "7" in question 48B.

48A) What is your height in feet?

Feet

48B) and inches?

Inches

49) What is your weight in pounds?

Pounds

50) What is your year in school?

Select One

51) What is your enrollment status?

○ Full-time   ○ Part-time   ○ Other

52) Have you transfered to this college or university within the last 12 months?

○ No   ○ Yes

53) How do you usually describe yourself? (Mark all that apply)

☐ White, non Hispanic (includes Middle Eastern)
☐ Black, non Hispanic
☐ Hispanic or Latino/a
☐ Asian or Pacific Islander
☐ American Indian, Alaskan Native, or Native Hawaiian
☐ Biracial or Multiracial
☐ Other

54) Are you an international student?

No   Yes

Revised version of the online American College Health Association-
National College Health Assessment (ACHA-NCHA)

○   ○

55) What is your relationship status?

○ Not in a relationship
○ In a relationship but not living together
○ In a relationship and living together

56) What is your marital status?

○ Single
○ Married/Partnered
○ Separated
○ Divorced
○ Other

57) Where do you currently live?

○ Campus residence hall
○ Fraternity or sorority house
○ Other college/university housing
○ Parent/guardian's home
○ Other off-campus housing
○ Other

58) Are you a member of a social fraternity or sorority? (e.g., National Interfraternity Conference, National Panhellenic Conference, National Pan-Hellenic Council, National Association of Latino Fraternal Organizations)

○ No ○ Yes

59) How many hours a week do you work for pay?

Select One

60) How many hours a week do you volunteer?

Select One

61) What is your primary source of health insurance?

○ My college/university sponsored plan
○ My parents' plan
○ Another plan
○ I don't have health insurance
○ I am not sure if I have health insurance

62) What is your approximate cumulative grade average?

Revised version of the online American College Health Association-
National College Health Assessment (ACHA-NCHA)

○ A   ○ B   ○ C   ○ D/F   ○ N/A

63) In the last 12 months, have your participated in organized college athletics at any of the following levels? (Please mark the appropriate column for each row)

	No	Yes
Varsity	○	○
Club sports	○	○
Intramurals	○	○

64) Do you have any of the following disabilities or medical conditions? (Please mark the appropriate column for each row)

	No	Yes
Attention Deficit and Hyperactivity Disorder (ADHD)	○	○
Chronic illness (e.g., cancer, diabetes auto-immune disorders)	○	○
Deaf/Hard of hearing	○	○
Learning disability	○	○
Mobility/Dexterity disability	○	○
Partially sighted/Blind	○	○
Psychiatric disability	○	○
Speech or language disorder	○	○
Other disability	○	○

Thank you for taking the time and thought to complete this survey. We appreciate your participation!

[Previous Page]   [Submit Survey]

Appendix C:
American College Health Association-National College Health Assessment
(ACHA-NCHA) Revisions Committee Members

ACHA MEMBERS	ACHA STAFF

ACHA MEMBERS

Sherri Darrow, Ph.D.
Director, Student Health and Wellness
Social and Preventive Medicine
University at Buffalo

Alan Glass, M.D.
Director, Student Health Services
Washington University in St. Louis

Michael Haines, M.S.
Former Director
National Social Norms Resource Center
Northern Illinois University

Patricia Ketcham, Ph.D., CHES
Associate Director, Health Promotion
Student Health Services
Oregon State University

James (Jim) Mallinson, M.A.
Director (retired), Student Health Center
University of North Carolina at
Charlotte

Alejandro Martinez, Ph.D.
Senior Associate Director
Vaden Student Health Services
Stanford University

Daisye Orr, MPH, CHES
Health Educator
Washington State Department of Health

Cheryl Presley, Ph.D.
Former Director
CORE Institute
Southern Illinois University

ACHA STAFF

Mary Hoban, Ph.D., CHES
Director, ACHA-NCHA Program Office

Theresa Jackson, MPH
Research Assistant, ACHA-NCHA
Program Office

E. Victor Leino, Ph.D.
Research Director, American College
Health Association

ACHA-NCHA Web Version
Frequently Asked Questions

CONTACTING STUDENTS

Q: How are the students contacted using the NCHA-Web?
A: ACHA contacts students to participate in the ACHA-NCHA through a letter of invitation/consent sent via blast e-mail. You provide the text of the letter of invitation to ACHA as well as the subject line you want included with the e-mail. The ACHA-NCHA Program Office staff is then responsible for programming your school's individual survey. After you decide who you wish to survey (e.g., all registered students, 10% of registered students), you must also supply ACHA with a list of student e-mail addresses in the first column of an Excel spreadsheet file. If you choose to sample only a percentage of your students, ACHA has the ability to select a random selection of students from a larger list (e.g., all student e-mails) and will do so for no additional charge. Additionally, you may want to personalize the electronic letter of invitation/consent sent to each student with the student's first name, such as "Dear Mary". If so, add the first name corresponding to each e-mail address in the second column of the spreadsheet file. This information can usually be found at your Registrar's Office or Office of Institutional Research. After receiving the Excel file with the e-mail addresses, ACHA generates a unique ID for each student. This unique ID number is imbedded in the survey URL sent to each students. Unique IDs are downloaded with student responses and allow us to prevent duplicate responses and contact non-responders. Unique IDs may also be used to distribute incentives to survey participants. Once survey data are downloaded, ACHA deletes Unique IDs from our records.

Q: Does ACHA send reminder messages to all students or just those who have not responded to the survey?
A: ACHA tracks survey responders via the unique ID. ACHA contacts only non-responders with a reminder e-mail. Although you may use your initial letter of invitation/consent and subject line to contact non-responders, many schools choose to compose a different letter of invitation/consent and use a different subject line when doing so.

Q: Is the NCHA-Web considered anonymous or confidential?
A: The web version of the survey is confidential. E-mail addresses are never physically attached to survey responses. The unique ID number is attached to student e-mail addresses solely for the purpose of contacting non-responders and awarding incentives. Once the surveying is completed, the spreadsheet file holding the e-mail addresses and unique IDs is erased. Unique IDs are also deleted from the downloaded survey data.

Q: What if my school will not allow me to provide ACHA with students' e-mail addresses?
A: If your campus makes student e-mail addresses public via a published student directory, they will likely approve e-mail access by ACHA to administer the ACHA-NCHA. If your campus prohibits the exchange of e-mail addresses, you may still administer the NCHA-Web without providing student e-mail addresses to ACHA. ACHA will assign your institution unique IDs based on the number of students you intend to sample. Then, your campus IT department will have to match e-mail addresses to unique IDs and write computer code that will provide a unique URL to each student based on the unique ID generated by ACHA. Thus, each student will still receive her/his own unique URL in the e-mail invitation sent from your school. You would then be responsible for contacting non-responders. ACHA can provide your school with a list of submitted unique IDs so they may be eliminated prior to re-contact.

Q: My campus will not allow ACHA access to student e-mail addresses and my IT department has told me that they can only blast e-mail "static messages" (identical messages). The e-mail messages cannot be personalized with a unique ID imbedded in the URL nor in the greetings. Is there another way to administer the ACHA-NCHA Web?

A. If there is no way to remedy this problem on campus, ACHA can provide a "static" URL for your campus survey submissions. However, without unique IDs imbedded in the URL, ACHA cannot control for multiple submissions, cannot track non-responders, and cannot identify respondents to award incentives. While this methodology has been implemented with success, it is not recommended as it introduces a possible source of error not present in the standard methodology.

Q: Do I have to use the ACHA letter of invitation/consent as provided in the User's Manual?
A: ACHA offers an example letter of invitation/consent in the User's Manual, which you may edit to meet the specific needs of your campus or requirements of your IRB. Whether or not you choose to use the example letter of invitation/consent, you must provide ACHA with the exact text of your letter of invitation/consent as well as the exact text you wish to use in reminder messages sent to non-responders. You must additionally provide a subject line for the e-mail containing the letter of invitation/consent or reminder letter of invitation/consent. The standard subject line used is "ACHA-National College Health Assessment (ACHA-NCHA)", but you may want to personalize the subject line to include your school name so your students know the communication is connected with your college/university.

Q: How do ACHA-NCHA-Web survey participants provide consent?
A: The example letter of invitation/consent includes instructions that stipulate students who are 18 years of age or older need only click on the link to the survey to confirm their consent. If there is additional consent information required by your campus IRB, you may want to include a link in your invitation e-mail that will direct students to further information, such as a campus consent requirement. It is recommended that you keep the letter of invitation/consent clear and concise.

SAMPLE SIZE

Q: We are a school with a student population of 2,400. How many students should we contact using the ACHA-NCHA-Web?
A: The ACHA-NCHA User's Manual presents recommended sample sizes, or you can see the table of sample sizes listed below. Most ACHA-NCHA Web surveyors experience a range of 10% to 35% return. The average response proportion is about 25%. Thus, you may want to estimate the number of contacts based on a 25% return. With a school your size you should survey all 2,400 students with the aim of getting 600 submitted surveys.

The net proportion of returned surveys depends in part on the following: (1) campus-wide pre-publicity, (2) timing of first contact (do not survey just before finals or breaks), (3) contact of non-responders, (4) accuracy and utilization of e-mail addresses (e.g., do the students use the campus generated e-mail addresses?), (5) coordination with campus IT department to prevent interference from spam and/or rate control filters, and (6) incentives.

Student Population	Desired Sample Size (Returned Surveys)
<600	all students
600-3,000	600
3,000-9,999	700
10,000-19,999	800
20,000-29,999	900
•30,000	1,000

Q. If I sample all students or an entire subset of students (such as all first-year students or all graduate students) will our data be included in the final national ACHA-NCHA Reference Group even though we didn't select a random sample?

A. The purpose of sampling a randomized subset of students is simply to eliminate the need to sample all students. Thus, sampling the universe is the same or better than selecting a random subset of students. By sampling the universe of all students your data will be included in the Reference Group data set and reports.

INCENTIVES

Q: Can ACHA help us select random respondents to award participation incentives?
A: ACHA tracks ID numbers of participating students. We are able to randomly select ID numbers from the pool of participating students and provide you with the corresponding e-mail address from the spreadsheet file. E-mail addresses are never physically or electronically attached to the survey responses.

EXTRA QUESTIONS

Q: Can we add extra questions to the ACHA-NCHA Web?
A: Yes, it is possible to add up to five additional questions for a surcharge of $700. Your Institutional SPSS Data Set will contain the additional variables (extra questions) and they will be summarized on your Institutional Report.

The extra questions must be single questions. When asking a general question with an option to "select all that apply" from a list of responses, each item on the list is individually counted as a question. For example, the following question is actually five additional questions on the ACHA-NCHA:

> In the last 12 months, did a medical professional ever prescribe you the following narcotics? (Select all that apply)
> a. Codeine
> b. Hydrocodone
> c. Oxycodone
> d. Tramadol
> e. Morphine

Individual questions may not have more than 10 response categories. A minimum of three weeks is required to design the survey additions.

To order more than five questions or any outside of the criteria outlined above, you must contact Mary Hoban at mhoban@acha.org, and attach a draft of the extra questions.

ADDITIONAL DATA DOWNLOADS

Q: I have access to additional information from the registrar's office about the students in my sample. Is it possible to have these data merged with my NCHA data set for each subject?
A: Yes, some customers choose to merge other variables provided by the registrar's office (e.g., major, school or college, student athlete status, student residence hall) into their NCHA data set. It is possible to automatically download up to five additional variables for into your campus NCHA data file. Simply include the additional variables beginning in the third column of the Excel file containing the student e-mail addresses. The additional variables must be in numeric form. You are responsible for coordinating and maintaining the codebook of variable values and numeric codes. For example, if you want to be able to analyze your data by major and the registrar's office provides you with the major for each student selected to participate, you are responsible for developing the code (e.g. electrical engineering = 1) and providing the resulting numeric code to ACHA in your Excel file. When we return your campus NCHA data set to you,

the numeric value for these additional variables will be added matched for each respondent and added to your campus data set.

All variables sent to ACHA for automatic download into your NCHA dataset must be described in your IRB application.

COSTS

Q: What is the cost of implementing the ACHA-NCHA Web?
A: The cost is less for the survey if your school is an ACHA Institutional Member — see the pricing chart at http://www.acha.org/projects_programs/NCHA_Pricing.pdf. You are charged for all initial contacts ($0.20 each), re-contacting non-responders ($0.10 each), and an additional processing fee for each survey that is submitted ($0.25 each).

ESTIMATED ACHA INSTITUTIONAL MEMBER PRICE
BASED ON 3,000 CONTACTS

	Price per contact	Number of contacts	Total cost
Initial contacts	$0.20	3,000 (100%)	$600.00
1ˢᵗ Contact Non-responders	$0.10	2,700 (90%)	$270.00
2ⁿᵈ Contact Non-responders	$0.10	2,400 (80%)	$240.00
Submitted Surveys	$0.25	750 (25%)	$187.50
Report Package		1	$300.00
TOTAL			$1,597.50

If your contact list is 3,000, you would pay for each initial contact ($600). First contact of non-responders would be about 90% of the 3,000 or 2,700 ($270). Second contact of non-responders would be about 80% of 3,000 or 2,400 ($240). Final submitted surveys would be about 25% of the original sample of 3,000 or 750 ($187.50). Report packages are $300 each for ACHA Institutional Members. The estimated total for this example of 3,000 students would be $1,597.50.

REQUIRED PAPERWORK

Q: Why do I need to fill out an Institution of Higher Education Demographic Survey?
A: All colleges and universities that participate in the ACHA-NCHA are required to complete a demographic survey. The Demographic Survey, which can be found at http://www.acha.org/projects_programs/ncha_participate.cfm#forms, contains questions specific to your school, such as number of students, campus location, Carnegie classification, and survey methodology. ACHA uses this information to report on the aggregate dataset from each survey period. As ACHA will not process your school's dataset or reports until this information is on file, it is recommended that you complete this survey as soon as possible once surveying has commenced.

Q: Am I required to obtain IRB (Institutional Review Board) approval for conducting the ACHA-NCHA?
A: Each surveyor is responsible for securing approval for conducting research involving human subjects from the appropriate campus office, prior to implementing the ACHA-NCHA. The body that grants approval for such surveys on your campus may be referred to as an Institutional (or Internal) Review Board, Human Subjects Committee, Grant Review Board, Research Review Board, or something similar. Many campuses will determine that the ACHA-NCHA presents minimal risk to the participants and grant an exemption or offer an expedited review of the

project. Other campuses will require a full review of your proposed research with the ACHA-NCHA. Whatever the case, you are responsible for securing approval for the ACHA-NCHA, any additional questions added by the campus, and any additional variables you are providing to ACHA for automatic download into your dataset. A copy of such approval must be provided to ACHA before your dataset and reports are processed.

Q: Our school does not have an IRB, how can we meet this requirement?
A: If your campus does not have an IRB, we will accept written permission to survey from a campus administrator (e.g., Vice President/Vice Chancellor, Dean of Students). ACHA must have this information on file before processing your school's dataset and report packages. Therefore, please ensure you have received the appropriate approval prior to inviting your students to participate in the ACHA-NCHA.

REPORT PACKAGE

Q: What does the Report Package include?
A: The ACHA-NCHA Report Package contains your Institutional Executive Summary (15 pages of salient results selected by the ACHA-NCHA Advisory Committee), Institutional Report (50 pages of frequency distributions for every variable in the survey for all subjects, and by each sex, with a bar graph), and your institutional data on a CD (SPSS system file with built in codebook). You will also receive a copy of the prior year's Reference Group Executive Summary so you may begin making comparisons with your institutional data as soon as you receive your reports. After the survey period is completed for all schools involved, you will receive the Reference Group Executive Summary and Reference Group Report for the survey period. Each report package is $300 for ACHA Institutional Members and $500 for ACHA nonmembers.

Q: Is it possible to get extra report packages for a subset of my students, and if so, how much does it cost?
A: Many institutions choose to have ACHA generate extra report packages for a subset of their students (e.g., graduates, undergraduates, residents, first-year students, students on only one of multiple campuses). Each additional Reports package is $300 ($150 each for the Executive Summary and Institutional Report). Please contact Mary Hoban at mhoban@acha.org if you are interested in purchasing extra reports for your school.

SPAM Filters and Rate Control Systems

Q: My school uses rate control systems and spam filters that can intercept and block mass e-mailings from outside organizations. What can we do to prevent our invitations to participate in ACHA-NCHA Web from being blocked?
A: It is important that you work closely with your IT department to try to prevent this problem. Spam filters often look for specific words in the subject line and body of a message in an effort to decrease unsolicited e-mail messages. Words such as "win," "prizes," "last chance," "drawing," or even the use of multiple "!!!!!" in a message can trigger a spam filter and prevent the distribution of the messages. It's important to test messages against your spam filter before the survey period begins. Rate control filters tend to block messages from a single IP address once a predetermined limit has been exceeded.

Please notify the IT department of your survey plans and let them know that that the invitations to participate will originate from webemailer.com (IP address 38.119.108.121) or webemailer.net (IP address 38.119.108.120). Although the e-mail messages will originate from one of the webemailer domains, they will appear to be from NCHA-WEB@acha.org (IP address

67.151.72.91). All three IP addresses and domains should be "whitelisted" to prevent the rate control system or spam filter from blocking our invitations. Likewise, the IT department may be able to "lift a quarantine" on e-mails originating from webemailer.com, webemailer.net, and acha.org. Lastly, the predetermined limit on rate control systems might need to be increased during your survey period.

SECURITY AND DOWNLOADING

Q. Does ACHA host the ACHA-NCHA Web internally on their servers and what kind of security is provided?
A. ACHA uses WebSurveyor Corporation servers to host the web version and contact students. WebSurveyor Corporation uses the most advanced encryption, firewalls, backup and 24-hour service support. As a major survey corporation, the security and access to information of any kind is under strict regulations and procedures that go beyond the measures taken by most organizations, agencies, and institutions of higher education. Additionally, ACHA does not provide any other organizations with students e-mail addresses.

Please see the ACHA-NCHA User's Manual as well the Security and Downloading Synopsis for more information at:
http://www.acha.org/projects_programs/ncha_websecurity.cfm

Q. My IRB wants to know what kind of training ACHA-NCHA staff has with regard to protection of human subjects in research.
A. All ACHA-NCHA staff have completed the Human Participants Protection Education for Research Teams online course, sponsored by the National Institutes of Health. Certificates of completion are on file at ACHA and available upon request.

MATERIALS NEEDED

Q: What do I need to send to ACHA to implement the NCHA-Web?
A: The following materials must be sent to ACHA at least two weeks prior to the desired survey launch date:
- The NCHA Order form, available at http://www.acha.org/projects_programs/ncha_participate.cfm#forms, and initial payments for both the total number of student contacts and report package.
- An Excel file of student e-mail addresses in the first column of the file, the student's first name in the second column (if the greetings are to be personalized), and any other variables to be downloaded (if applicable).
- The letter of invitation/consent and subject line.
- The reminder letter of invitation/consent that is sent to non-responders. To do this version, simply add to the initial invitation letter a phrase such as, "Recently you were contacted to participate in the ACHA-NCHA. Please consider…"

Please note that if you choose to add additional questions to the ACHA-NCHA, a final list of questions and all materials listed above much be sent to ACHA at least three weeks prior to the desired survey date.

Q: What do I need to send ACHA to process the ACHA-NCHA Web and receive my reports?
A:
- Your IRB or Administrative approval to survey (including approval of any additional questions or downloaded variables).

American College Health Association-National College Health Assessment
Frequently Asked Questions Document

- The Demographic Survey
  (www.acha.org/projects_programs/ACHA_NCHA_DEMOGRAPHIC_SURVEY.pdf)

ADDITIONAL RESOURCES
- ACHA-NCHA Website (www.acha-ncha.org)
- ACHA-NCHA Brochure
- ACHA-NCHA Users Manual
- Mary T Hoban, PhD, CHES
  Director, ACHA-NCHA Program Office
  410-859-1500 x216
  mhoban@acha.org

UNIVERSITY OF
MARYLAND
INSTITUTIONAL REVIEW BOARD

2100 Blair Lee Building
College Park, Maryland 20742-5121
301.405.4212 TEL 301.314.1475 FAX
irb@deans.umd.edu
www.umresearch.umd.edu/IRB

February 1, 2008

MEMORANDUM
Application Approval Notification

To:	Elbert Glover, Ph. D. Theresa Jackson Hughes, MPH Department of Public & Community Health
From:	Roslyn Edson, M.S., CIP IRB Manager University of Maryland, College Park
Re:	**IRB Application Number: # 08-0022** **Project Title:** "Examining the reliability and validity of mental health indicators on a revised national survey measuring college student health: Results of an experimental field pre-test"
Approval Date:	**January 24, 2008**
Expiration Date:	**January 24, 2009**
Type of Application:	Initial
Type of Research:	Non-Exempt
Type of Review For Application:	Expedited

The University of Maryland, College Park Institutional Review Board (IRB) approved your IRB application. The research was approved in accordance with 45 CFR 46, the Federal Policy for the Protection of Human Subjects, and the University's IRB policies and procedures. Please reference the above-cited IRB application number in any future communications with our office regarding this research.

Recruitment/Consent: For research requiring written informed consent, the IRB-approved and stamped informed consent document is enclosed. The IRB approval expiration date has been stamped on the informed consent document. Please keep copies of the consent forms used for this research for three years after the completion of the research.

Continuing Review: If you intend to continue to collect data from human subjects or to analyze private, identifiable data collected from human subjects, after the expiration date for this approval (indicated above), you must submit a renewal application to the IRB Office at least 30 days before the approval expiration date.

Modifications: Any changes to the approved protocol must be approved by the IRB before the change is implemented, except when a change is necessary to eliminate apparent immediate hazards to the subjects. If you would like to modify the approved protocol, please submit an addendum request to the IRB Office. The instructions for submitting a request are posted on the IRB web site at:
http://www.umresearch.umd.edu/IRB/irb_Addendum%20Protocol.htm.

Unanticipated Problems Involving Risks: You must promptly report any unanticipated problems involving risks to subjects or others to the IRB Manager at 301-405-0678 or redson@umresearch.umd.edu.

Student Researchers: Unless otherwise requested, this IRB approval document was sent to the Principal Investigator (PI). The PI should pass on the approval document or a copy to the student researchers. This IRB approval document may be a requirement for student researchers applying for graduation. The IRB may not be able to provide copies of the approval documents if several years have passed since the date of the original approval.

Additional Information: Please contact the IRB Office at 301-405-4212 if you have any IRB-related questions or concerns.

REFERENCES

Aday, L. A. & Cornelius, L. J. (2006). *Designing and conducting health surveys: A comprehensive guide* (3rd edition). San Francisco: Jossey-Bass.

Aiken, L. R. (1997). *Psychological testing and assessment*. Boston: Allyn and Bacon.

American Association for Higher Education. (1991). *Nine principles of good practice for assessing student learning.* Sterling, VA: Stylus.

American College Health Association [ACHA]. (2007a). American College Health Association National College Health Assessment Spring 2006 Reference Group Data Report (Abridged). *Journal of American College Health, 55*, 196-206.

American College Health Association [ACHA]. (2007b). American College Health Association-National College Health Assessment. Retrieved December 1, 2007 from http://www.acha-ncha.org/.

American College Health Association [ACHA]. (2007c). American College Health Association: Who we are. Retrieved February 15, 2008 from http://www.acha.org/about_acha/ACHA_WhoWeAre.pdf.

American College Health Association [ACHA]. (2006a). American College Health Association National College Health Assessment Spring 2004 Reference Group Data Report (Abridged). *Journal of American College Health, 54*, 201-211.

American College Health Association [ACHA]. (2006b). American College Health Association National College Health Assessment Spring 2005 Reference Group Data Report (Abridged). *Journal of American College Health, 55*, 5-16.

American College Health Association [ACHA]. (2005). American College Health Association National College Health Assessment Spring 2003 Reference Group Data Report. *Journal of American College Health, 53*, 199-210.

American College Health Association [ACHA]. (2004a). *National College Health Assessment: Reliability and Validity Analyses 2000*. Baltimore, MD: Author.

American College Health Association [ACHA]. (2004b). ACHA-NCHA user's manual. Retrieved February 14, 2008 from the ACHA-NCHA website at http://www.acha-ncha.org/docs/ACHA-NCHA_USERS_MANUAL_ver2-004.pdf.

Anxiety Disorders Association of America [ADAA]. (2007). An audit of mental health care at U.S. colleges and universities: Focus on anxiety disorders. Retrieved January 14, 2008 from http://www.adaa.org/Bookstore/Publications/ADAA%20Report%20-%20FINAL%20embargoed.pdf.

Appelbaum, P. S. (2006). "Depressed? Get out!": Dealing with suicidal students on college campuses. *Psychiatric Services, 57*(7), 914-916

Areni, C. S., Ferrell, M. E., & Wilcox, J. B. (1999). The effects of need for cognition and topic importance on the latency and extremity of responses to attitudinal inquiry. *Advances in Consumer Research, 26*(1), 63-68.

Azar, B. (2007). A case for angry men and happy women. *Monitor on Psychology, 38*(4), p. 18.

Baldwin, W. (2000). Information no one else knows: The value of self-report. In A.A. Stone, J. S. Turkkan, C. A. Bachrach, J. B. Jobe, H. S. Kurtzman, & V. S. Cain (Eds). *The Science of self-report: Implications for research and practice*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Barker, P., Gfroerer, J., Casper, R., & Lessler, J. (1998). Major design changes in the National Household Survey on Drug Abuse. *Proceedings of the Survey Research Methods Section of the American Statistical Association*, *25*, 732-737.

Baum, S. & Payea, K. (2005). *Education pays 2004: The benefits of higher education for individuals and society [revised]*. Retrieved December 15, 2007 from the College Board website at http://www.collegeboard.com/prod_downloads/ press/cost04/ EducationPays2004.pdf.

Beeber, L. S. (1999). Testing an explanatory model of the development of depressive symptoms in young women during a life transition. *Journal of American College Health, 47*, 227-234.

Benton, S. A., Robertson, J. M., Tseng, W. C., Newton, F. B., & Benton, S. L. (2003). Changes in counseling center client problems across thirteen years. *Professional Psychology: Research and Practice 34, 66-72.*

Bhave, S. & Nagpal, J. (2005). Anxiety and depressive disorders in college youth. *Pediatric Clinics of North America, 52*(1), 97-134.

Blom, S., & Beckley, S. (2005, January 28). 6 major challenges facing student health programs. *Chronicle of Higher Education, 51*(21), B25-B26.

Blumenthal, S. J., & Kepfer, D. J. (1990). Suicide over the life cycle: Risk factors. In *Assessment and Treatment of Suicidal Patients*. Washington, D.C.: American Psychiatric Press, Inc., 322-337.

Bowden, A., Fox-Rushby, J. A., Nyandieka, L., & Wanjau, J. (2002). *Health Policy and Planning, 17*(3), 322-330.

Braun, D. L., Sunday, S. R., Huang, A., & Halmi, K. A. (1999). More males seek treatment for eating disorders. *International Journal of Eating Disorders, 25*(4), 415-424.

Brindis, C., & Reyes, P. (1997). At the crossroads: Options for financing college health services in the 21st century. *Journal of American College Health, 45*(6), 279-288.

Brinson, J. A. & Kottler, J. A. (1995). Minorities' underutilization of counseling centers' mental health services: A case for outreach and consultation. *Journal of Mental Health Counseling, 17*, 371-386.

Brown, S. A., Williams, K., & Collins, A. (2007). Past and recent deliberate self-harm: Emotion and coping strategy differences. *Journal of Clinical Psychology, 63*(9), 791–803.

Carmine, E. G. & Zeller, R. A. (1979). *Reliability and validity assessment*. Beverly Hills: Sage.

Caulfield, S. C. (2001). Depression on college campuses: A new challenge for student health services and counseling centers. *Student Health Spectrum*, *Fall*, 18–11.

Centers for Disease Control and Prevention [CDC]. (2007). Health and safety for college students. Retrieved January 5, 2008 from http://www.cdc.gov/Features/ CollegeHealth/.

Centers for Disease Control and Prevention [CDC]. (1997). Youth Risk Behavior Surveillance: National College Health Risk Behavior Survey – United States, 1995. *MMWR, 46*(SS-6), 1-54.

The College Board. (2007). Trends in College Pricing 2007. Retrieved December 15, 2007 from http://www.collegeboard.com/prod_downloads/ about/news_info/trends/trends_pricing_07.pdf.

Collins, M. E. & Mowbray, C. T. (2005). Higher education and psychiatric disabilities: National survey of campus disability services. *American Journal of Orthopsychiatry, 75*(2), pp. 304–315.

Conrad, F. G. & Blair, J. (2004). Data quality in cognitive interviews: The case of verbal reports. In S. Presser, J. M. Rothgeb, M. P. Couper., et al. (eds.), *Methods for testing and evaluating survey questionnaires*. Hoboken, NJ: John Wiley & Sons, pp. 67-87.

Cook, L.J. (2007). Striving to help college students with mental health issues. *Journal of Psychosocial Nursing, 45*(4), 40-44.

The Core Institute (2007). *Alcohol and drug surveys*. Retrieved January 20, 2008 from the Core Institute website at: http://www.siu.edu/~coreinst/.

Couper, M. P. (2007). Issues of representation in eHealth research (with a focus on web surveys). *American Journal of Preventive Medicine, 32*(5S), S83-S89.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297-334.

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

DeMaio, T., Rothgeb, J., & Hess, J. (1998). Improving survey quality through pretesting. *Statistical Research Division Working Papers in Survey Methodology, #98-03*. Washington, DC: U.S. Bureau of the Census. Retrieved January 5, 2008 from http://www.census.gov/srd/papers/pdf/sm98-03.pdf.

De Vaus, D. (1995). *Surveys in social research* (4th edition). London: Routledge.

DeVellis, R. F. (2003). *Scale development: Theory and applications* (2nd edition). Thousand Oaks: Sage.

Di Iorio, C. K. (2005). *Measurement in health behavior: Methods for research and education*. San Francisco: Jossey-Bass.

Dillman, D. (2000). *Mail and internet surveys: The tailored design method* (2nd edition). New York: John Wiley and Sons.

Douglas, K. A., Collins, J. L., Warren, C., Kann, L., Gold, R., Clayton, S., Ross, J. G., & Kolbe, L. J. (1997). Results from the 1995 National College Health Risk Behavior Survey. *Journal of American College Health, 46*(2), 55-66.

Drewnowski, A., Yee, D. K., Kurth, C. L., & Krahn, D. D. (1994). Eating pathology and *DSM—III—R* bulimia nervosa: A continuum of behavior. *American Journal of Psychiatry, 151*, 1217-1219.

Dunteman, G. H. (1989). *Principal components analysis*. Newbury Park, CA: Sage.

Epstein, J. & Klinkenberg, W. D. (2001). From Eliza to Internet: A brief history of computerized assessment. *Computers in Human Behavior, 17*, 295-314.

Eysenbach, G. & Wyatt, J. (2002). Using the internet for surveys and health research. *Journal of Medical Internet Research, 4*(2), e13.

Field, A.P. (2005). *Discovering statistics using SPSS* (2nd edition). London: Sage.

Fowler, F. J. (2004). The case for more split-sample experiments in developing survey instruments. In S. Presser, J. M. Rothgeb, M. P. Couper., et al. (Eds). *Methods for testing and evaluating survey questionnaires*. Hoboken, NJ: John Wiley & Sons, pp.173-188.

Fowler, F. J. (1995*). Improving survey questions: Design and evaluation*. Thousand Oaks, CA: Sage.

Fowler, F. J. (2002). *Survey research methods* (3rd edition). Thousand Oaks, CA: Sage.

Gallagher, R. P. (2007). *National Survey of Counseling Center Directors*. Retrieved January 2, 2008 from the International Association of Counseling Services, Inc. website: http://www.iacsinc.org/NsccdSurveyFinal_v2.pdf.

Garson. (n.d.). *Chi-square significance tests*. Retrieved February 10, 2008 from the North Carolina State University website at: http://www2.chass.ncsu.edu/garson/ pA765/chisq.htm.

Geddes, J., Freemantle, N., Harrison, P., & Bebbington, P. (2000). Atypical antipsychotics in the treatment of schizophrenia: A systematic overview and meta-regression analysis. *British Medical Journal, 321,* 1371–1376.

Gilbert, S. P. (1992). Ethical issues in the treatment of severe psychopathology in university and college counseling centers. *Journal of Counseling & Development, 70*, 695-699.

Goldsmith, S. K., Pellmar, T. C., Kleinman, A. M., & Bunney, W. E. (Eds.). (2002). *Reducing suicide: A national imperative*. Washington, DC: The National Academies Press.

Grace, T. W. (1997). Health problems of college students. *Journal of American College Health, 45*(6), 243-250.

Gratz, K. L., Conrad, S. D., & Roemer, L. (2002). Risk factors for deliberate self-harm among college students. *American Journal of Orthopsychiatry, 72*, 128–140.

Hensher, D. A., Rose, J. M., & Greene, W. H. (2005). *Applied choice analysis: A primer*. Cambridge: Cambridge University Press.

Herrero, J. & Meneses, J. (2006). Short web-based versions of the perceived stress (PSS) and Center for Epidemiological Studies-Depression (CESD) Scales: A comparison to pencil and paper responses among internet users. *Computers in Human Behavior, 22*, 830-846.

Higher Education Research Institute [HERI]. (n.d.). *CIRP Freshman Survey: Reliability and validity*. Retrieved January 5, 2008 from http://www.gseis.ucla.edu/ heri/PDFs/CIRP_Reliability_Validity.PDF.

Hingson, R., Heeren, T., Winter, M., & Wechsler, H. (2005). Magnitude of alcohol-related mortality and morbidity among U.S. college students ages 18-24: Changes from 1998 to 2001. *Annual Review of Public Health, 26*, 259-279.

Hingson, R. W. & Howland, J. (2002). Comprehensive community interventions to promote health: Implications for college-age drinking problems. *Journal of Studies on Alcohol Supplement, 14*, 226-240.

Hoek, H. W. & van Hoeken, D. (2003). Review of the prevalence and incidence of eating disorders. *International Journal of Eating Disorders, 34*(4), 383-396.

Hudson, J. I., Hiripl, E., Pope, Jr., H. G., & Kessler, R. C. (2007). The prevalence and correlates of eating disorders in the National Comorbidity Survey Replication. *Biological Psychiatry, 61*(3), 348-358.

Institute for Higher Education Policy (1998). Reaping the benefits: Defining the public and private value of going to college. *The New Millennium Project on Higher Education Costs, Pricing, and Productivity*. Washington, DC: Author.

Johnston, L. D., O'Malley, P. M., Bachman, J. G., & Schulenberg, J. E. (2006). *Monitoring the Future national survey results on drug use, 1975–2005: Volume II, College students and adults ages 19–45* (NIH Publication No. 06-5884). Bethesda, MD: National Institute on Drug Abuse.

Joiner, T. E. & Blalock, J. A. (1995). Gender differences in depression: The role of anxiety and generalized negative affect. *Sex Roles, 33*(1-2), 91-108.

Jones, S. & Madden, M. (2002). The Internet goes to college: How students are living in the future with today's technology. *Pew Internet and American Life Project*. Retrieved December 5, 2007 from http://www.pewinternet.org.

Kadison, R. (2006). College psychiatry 2006: Challenges and opportunities. *Journal of American College Health, 54*(6), 338-340.

Kadison, R. & DiGeronimo, T. F. (2004). *College of the overwhelmed: The campus mental health crisis and what to do about it*. San Francisco: Jossey-Bass.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20*, 141-151.

Karras, B. T. & Tufano, J. T. (2006). Multidisciplinary eHealth survey evaluation methods. *Evaluation and Program Planning, 29*, 413-418.

Kessler, R. C. (1994). The National Comorbidity Survey of the United States. *International Review of Psychiatry, 6*, 365-376.

Kessler, R. C., Chiu, W. T., & Walters, E. (2005). Prevalence, severity, and comorbidity of 12-month *DSM-IV* disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62*, 617-627.

Kessler, R. C., Foster, C. L., Saunders, W. B., & Stang, P. E. (1995). Social consequences of psychiatric disorders: I. Educational attainment. *American Journal of Psychiatry, 152,* 1026–1032.

Kessler, R. C., McGonagle, K. A., Zhao, S., Nelson, C. B., Hughes, M., Eshleman, S., Wittchen, H. U., & Kendler, K. S. (1994). Lifetime and 12-month prevalence of DSMIII-R psychiatric disorders in the United States: Results from the National Comorbidity Study. *Archives of General Psychiatry*, *51*, 8-19.

Kisch, J., Leino, E. V., & Silverman, M. M. (2005). Aspects of suicidal behavior, depression, and treatment in college students. *Suicide and Life-Threatening Behavior, 35*(1), 3-13.

Kitzrow, M. A. (2003). The mental health needs of today's college students: Challenges and recommendations. *NASPA Journal, 41*(1), 167-181.

Knapp, L. G., Kelly-Reid, J. E., Whitmore, R. W., and Miller, E. (2007). *Enrollment in Postsecondary Institutions, Fall 2005; Graduation Rates, 1999 and 2002 Cohorts; and Financial Statistics, Fiscal Year 2005* (NCES 2007-154). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved December 15, 2007 from http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2007154.

Knight, J. R., Wechsler, H., Kuo, M., Seibring, M., Weitzman, E. R., & Schuckit, M. Alcohol abuse and dependence among U.S. college students. *Journal of Studies on Alcohol*, *63*(3), 263-270.

Korhonen, P. & Siljamaki, A. (1998). Ordinal principal component analysis. *Computational Statistics and Data Analysis, 26*, 411-424.

Kurth, C. L., Krahn, D. D., Nairn, K., & Drewnowski, A. (1995). The severity of dieting and bingeing behaviors in college women: Interview validation of survey data. *Journal of Psychiatric Research, 29*(3), 211-225.

Kypri, K., Gallagher, S. J., & Cashell-Smith, M. L. (2004). An Internet-based survey method for college student drinking research. *Drug and Alcohol Dependence, 76*, 45-53.

Lake, P. & Tribbensee, N. (2002). The emerging crisis of college student suicide: Law and policy responses to the serious forms of self-inflicted injury. *Stetson Law Review, 32*, 125-157.

Leino, E. V., & Kisch, J. (2005). Correlates and predictors of depression in college students. *American Journal of Health Education, 36*(2), 66-74.

Levine, A., & Cureton, S. (1998a). *When hope and fear collide: A portrait of today's college student.* San Francisco: Jossey Bass.

Levine, A., & Cureton, J. S. (1998b). What we know about today's college students. *About Campus, 3*(1), 4–9.

Marsh, C. (1982). *The Survey Method: The contribution of surveys to sociological explanation*. London: George Allen & Unwin Ltd.

Martin, E. (2004). Vignettes and respondent debriefing for questionnaire design and evaluation. In S. Presser, J. M. Rothgeb, M. P. Couper., et al. (eds.), *Methods for testing and evaluating survey questionnaires*. Hoboken, NJ: John Wiley & Sons, pp. 149-171.

McCarty, C. A., Ebel, B. E., Garrison, M. M., DiGiuseppe, D. L., Christakis, D. A., & Rivara, F. P. (2004). Continuity of binge and harmful drinking from late adolescence to early adulthood. *Pediatrics, 114*(3), 714-719.

McMahon, W. W. (1999). *Education and development: Measuring the social benefits*. New York: Oxford University Press.

Meilman, P. W., & Pattis, J. A. (1994). Suicide attempts and threats on one college campus: Policy and practice. *Journal of American College Health, 42*(4), 147-154.

Merikangas, K. R., Ames, M., Lihong, C., Stang, P. E., Ustun, T. B., Von Korff, M., Kessler, R. C. (2007). The impact of comorbidity of mental and physical conditions on role disability in the U.S. adult household population. *Archives of General Psychiatry, 64*, 1180-1188.

Miles, J. & Gilbert, P. (2005). *A handbook of research methods for clinical and health psychology*. United Kingdom: Oxford University Press.

Moore, J., Pascale, J., Doyle, P., Chan, A., & Griffiths, J. K. (2004). Using field experiments to improve instrument design: The SIPP methods panel project. In S. Presser, J. M. Rothgeb, M. P. Couper., et al. (Eds). *Methods for testing and evaluating survey questionnaires*. Hoboken, NJ: John Wiley & Sons, pp. 189-208.

Mowbray, C. T., Megivern, D., Mandiberg, J. M., et al. (2006). Campus mental health services: Recommendations for change. *American Journal of Orthopsychiatry, 76*(2), 226–237.

National Eating Disorders Association [NEDA]. (2005). Statistics: Eating disorders and their precursors. Retrieved January 20, 2008 from http://www.sc.edu/healthycarolina/pdf/facstaffstu/eatingdisorders/EatingDisorderStatistics.pdf.

National Institute of Alcoholism and Alcohol Abuse [NIAAA]. (2002). *A call to action: Changing the culture of drinking at U.S. colleges.* Retrieved January 5, 2008 from the NIAAA website: http://www.collegedrinkingprevention.gov/media/TaskForceReport.pdf.

National Institute of Mental Health [NIMH]. (2007a). *The numbers count: Mental disorders in America*. Retrieved January 6, 2008 from http://www.nimh.nih.gov/health/publications/the-numbers-count-mental-disorders-in-america.shtml.

National Institute of Mental Health [NIMH]. (2007b). *Anxiety disorders*. NIH Publication No. 06-3879. Retrieved January 14, 2008 from http://www.nimh.nih.gov/health/publications/anxiety-disorders/nimhanxiety.pdf.

National Institute of Mental Health [NIMH]. (2007c). *Depression.* NIH publication No. 07-3561. Retrieved January 15, 2008 from http://www.nimh.nih.gov/health/publications/depression/nimhdepression.pdf.

National Institute of Mental Health [NIMH]. (2007d). *Eating disorders*. NIH Publication No. 07-4901. Retrieved January 15, 2008 from http://www.nimh.nih.gov/health/publications/eating-disorders/nimheatingdisorders.pdf.

National Mental Health Association & The Jed Foundation [NMHA & Jed]. (2002).

   Safeguarding your students against suicide: Expanding the safety network.

   Alexandria, VA: Author.

Nunnally, J. C. & Berstein, J. H. (1994). *Psychometric theory*. New York: McGraw-Hill.

Oswalt, R., & Finkelberg, S. (1995). College depression: Causes, duration, and coping.

   Psychological Reports, 77, 858.

Ouimet, J., Bunnage, B., Carini, R. M., Kuh, G. D., & Kennedy, J. (2004). Using focus

   groups to establish the validity and reliability of a college student survey.

   Research in Higher Education, 45, 233-250.

Padron, J. M. (2006). Experience with post-secondary education for individuals with

   severe mental illness. *Psychiatric Rehabilitation Journal, 30*(2), 147-149.

Pagano, M. & Gauvreau, K. (2000). *Principles of biostatistics* (2nd edition). Pacific

   Grove, CA: Duxbury/Thomson Learning.

Parks, K. A., Pardi, A. M., & Bradizza, C. M. (2006). Collecting data on alcohol use and

   alcohol-related victimization: A comparison of telephone and Web-based survey

   methods. *Journal of Studies on Alcohol, 67*, 318–323.

Pascarella, E.. T. (1998). Studying college students in the 21st century: Meeting new

   challenges. *The Review of Higher Education, 21*(2), 151-165. Retrieved April 7,

   2008 from https://muse.jhu.edu/journals/review_of_higher_education/

   v021/21.2pascarella.html.

Pealer, L. N. & Weiler, R. M. (2000). Web-based health survey research: A primer.

   American Journal of Health Behavior, 24, 69-72.

Pealer, L. N., Weiler, R. M, Pigg, R. M., Miller, D., & Dorman, S. M. (2001). The feasibility of a web-based surveillance system to collect health risk behavior data from college students. *Health Education & Behavior, 28*(5), 547-559.

Phillips. E. & Pratt, H. (2005). Eating disorders in college. *Pediatric Clinics of North America, 52*(1), 85-96.

Presley, C. A., Leichliter, M. A., & Meilman, P. W. (1998). *Alcohol and drugs on American college campuses: A report to college presidents: Third in a series, 1995, 1996, 1997.* Carbondale, IL: Core Institute, Southern Illinois University.

Presley, C. A., Meilman, P. W., & Lyerla, R. (1994). Development of the Core Alcohol and Drug Survey: Initial findings and future directions. *Journal of American College Health, 42*(6), 248-255.

Presser, S. & Blair, J. (1994). Survey pretesting: Do different methods produce different results? *Sociological Methodology, 24*, 73-104.

Presser, S., Couper, M. P., Lessler, J. T., Martin, E., Martin, J., Rothgeb, J. M., & Singer, E. (2004). Methods for testing and evaluating survey questions. *Public Opinion Quarterly*, *68*(1), 109-130.

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

Ross, S. E., Niebling, B. C., & Heckert, T. M. (1999). Sources of stress among college students. *College Student Journal, 33*(2), 312-317.

Rothgeb, J., Willis, G., & Forsyth, B. (2001). Questionnaire pretesting methods: Do different techniques and different organizations produce similar results? *Proceedings of the Annual Meeting of the American Statistical Association,* August 5-9, 2001.

Rowley, L. L., & Hurtado, S. (2002). *The non-monetary benefits of an undergraduate education*. University of Michigan: Center for the Study of Higher and Postsecondary Education.

Sax, L. J., Lindholm. J. A., Astin, A. W. Korn, W. S., & Mahoney, K. M. (2001). *The American freshman: National norms for fall 2001*. Higher Education Research Institute. University of California, Los Angeles.

Schwitzer, A. M., Bergholz, K., Dore, T., & Salimi, L. (1998). Eating disorders among college women: Prevention, education, and treatment responses. *Journal of American College Health, 46*(5), 199-207.

Sharkin, B.S. (2006). *College students in distress: A resource guide for faculty, staff, and campus community*. New York: Hasworth.

Shavelson, R. J. & Huang, L. (2003). Responding responsibly to the frenzy to assess learning in higher education. *Change, 35* (1), 10-19.

Shea, R. H. (2002). On the edge on campus. *U.S. News and World Report, 132*(5), 56-57.

Shisslak, C. M., Crago, M., & Estes, L. S. (1995). The spectrum of eating disturbances. *International Journal of Eating Disorders, 18*(3), 209-219.

Silverman, M., Meyer. P., Sloane, F., Raffel, M., & Pratt, D. (1997). The Big Ten student suicide study. *Suicide and Life Threatening Behavior, 27*, 285–303.

Singleton, R. A. & Straits, B. C. (1999). *Approaches to social research* (3rd edition). New York: Oxford University Press.

Smith, L. (2002). A tutorial on principal components analysis. Retrieved February 10, 2008 from the Agilent Technologies website at: http://www.chem.agilent.com/ cag/bsp/products/gsgx/Downloads/pdf/pca.pdf.

Soet, J. & Sevig, T. (2006). Mental health issues facing a diverse sample of college students: Results from the College Student Mental Health Survey. *NASPA Journal, 43*(3), 410-431.

Sontag, D. (2002). Who was responsible for Elizabeth Shin? *The New York Times, April 28.*

Sparling P. B., Snow T. K. (2002). Physical activity patterns in recent college alumni. *Research Quarterly for Exercise and Sport, 73*(2), 200-205.

Spector, P. E. (1992). *Summated rating scale construction: An introduction*. Thousand Oaks, CA: Sage.

Stevens, S. S. (1959). Measurement, psychophysics and utility. In C.W. Churchman & P. Ratoosh (Eds.), *Measurement definitions and theories*. New York: Wiley, pp. 18-63.

Substance Abuse and Mental Health Services Administration [SAMHSA]. (2007). *National registry of evidence-based programs and practices*. Retrieved January 5, 2008 from http://nrepp.samhsa.gov/about-evidence.htm.

Suicide Prevention Resource Center (SPRC). (2004). *Promoting mental health and preventing suicide in college and university settings*. Newton, MA: Education Development Center, Inc. Retrieved January 2, 2008 from the SPRC website: http://www.sprc.org/library/college_sp_whitepaper.pdf.

Sullivan, P. F. (1995). Mortality in anorexia nervosa. *American Journal of Psychiatry, 152*(7), 1073-1074

Suskie, L. (2006). Accountability and quality improvement. In P. Hernon, R. E. Dugan, & C. Schwartz (Eds.), *Revisiting outcomes assessment in higher education.* Westport, CT: Libraries Unlimited.

Thatch, L. (1995). Using electronic mail to conduct survey research. *Educational Technology, March-April*, 27-31.

Tourangeau, R. (1984). Cognitive science and survey methods: A cognitive perspective, in T. Jabine, M. Straf, J. Tanur, and R. Tourangeau (eds.), *Cognitive aspects of survey methodology: Building a bridge between disciplines.* Washington, DC: National Academy Press.

Tourangeau, R. (2004). Experimental design considerations for testing and evaluating questionnaires. In S. Presser, J. M. Rothgeb, M. P. Couper., et al. (eds.), *Methods for testing and evaluating survey questionnaires*. Hoboken, NJ: John Wiley & Sons, pp. 209-224.

Tourangeau, R., Rips, L., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge, MA: Cambridge University Press.

Trochim, W. M. K. (2006). *Research methods knowledge base*. Retrieved December 15, 2007 from http://www.socialresearchmethods.net/kb/.

U. S. Census Bureau. (2006). *School enrollment: Current population survey reports*. Retrieved February 1, 2008 from the U.S. Census Bureau website at: http://www.census.gov/population/www/socdemo/school.html.

U. S. Department of Health and Human Services [USDHHS]. (1999). *Mental health: A report of the Surgeon General*. Rockville, MD: U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Mental Health Services, National Institutes of Health, National Institute of Mental Health.

U. S. Department of Education. (2007). Office of Safe and Drug-Free Schools, Higher Education Center for Alcohol and Other Drug Abuse and Violence Prevention, *Experiences in effective prevention: The U.S. Department of Education's alcohol and other drug prevention models on college campuses grants,* Washington, D.C.

Voelker, R. (2007). Campus tragedy prompts closer look at mental health of college students. *Journal of the American Medical Association, 297*, 2335-2337.

Weale, M. (1992). The benefits of higher education: A comparison of universities and polytechnics. *Oxford Review of Economic Policy, 8*, 35-47.

Wechsler H., Lee J.E., Kuo M., Seibring M., Nelson T. F., & Lee H. P. (2002). Trends in college binge drinking during a period of increased prevention efforts: Findings from four Harvard School of Public Health study surveys, 1993-2001. *Journal of American College Health* 50(5):203-217.

Whitaker, L.C. (1989). Suicide and other crises. In: Grayson PA, Cauley K, (eds.), *College psychotherapy.* New York: Guilford, pp. 48-70.

Whitehead, L. C. (2007). Methodological and ethical issues in internet-mediated research in the field of health: An integrated review of the literature. *Social Science and Medicine, 65*, 782-791.

Whitlock, J., Eckenrode, J., & Silverman, D. (2006). Self-injurious behaviors in a college population. *Pediatrics, 117*, 1939–1948.

Woodside, D. & Garfinkel, P. (1992). Age of onset of eating disorders. *International Journal of Eating Disorders, 12*, 31-36.

Wright, D. B., Gaskell, G. D., & O'Muircheartaigh, C. A. (1998). Flashbulb memory assumptions: Using national surveys to explore phenomena. *British Journal of Psychology, 89*(Part 1), 103-121.