2005 AIR/NPEC RESEARCH GRANT PROPOSAL

Project Title: **A Longitudinal Study of Student Success: The Relation Between Academic Major, Student Demographics and Broad Student Outcomes**

Grant Amount Requested: **$ 29,822**

Principal Investigator:
Rachelle L. Brooks
Associate Director
Center for Assessment of Higher Education (CAHE)
1230 Benjamin Building
University of Maryland, College Park, MD 20742
Phone: **301.405.0778** Fax: **301.314.9207**
Email: **rlbrooks@umd.edu**

Co-Principal Investigator:
Dennis M. Kivlighan, Jr.
Chair, Dept. of Counseling and Personnel Services
3214 Benjamin Building
University of Maryland, College Park, MD 20742
Phone: **301.405.2863** Fax: **301.405.9995**
Email: **dennisk@umd.edu**

Co-Principal Investigator:
Janet A. Schmidt
Director of Interdisciplinary Research
Office of the Dean, College of Education
0108D Cole Student Activities Building
University of Maryland, College Park, MD 20742
Phone: **301.405.2347** Fax: **301.405.7138**
Email: **jschmidt@umd.edu**

Authorized Institutional Representative
Antoinette Lawson
Associate Director
Office of Research Administration and Advancement
3112 Lee Building
University of Maryland, College Park, MD 20742
Phone: **301.405.6269** Fax: **301.314.9569**
Email: **oraa@umd.edu**

_____ _____

Principal Investigator Co-PI

_____ _____

Co-PI Authorized Institutional Representative

PROJECT SUMMARY

This study addresses the assessment of a university's unique contribution to the development of broad student outcomes, such as thinking and reasoning abilities, by testing alternative measures both rooted within and spanning across academic disciplines. While the goal of critical thinking is common to most of higher education in the United States, institutions possess few tools and virtually no data to assess the extent to which graduating students demonstrate advancement in this area. An important step in the creation of an assessment is to determine how performance on discipline-specific outcome measures differs from that on items which cross disciplinary boundaries. Thus, one objective of the study will be to ground critical thinking questions within disciplinary groupings (e.g., social science, engineering, the arts), and to match them with students' major field concentrations. A second objective is to investigate the role of individual level characteristics (e.g. racial/ethnic status, socio-economic status, curricular/extracurricular activities, commuter status, etc.) in the achievement of critical thinking.

This study proposes a multi-method strategy which will utilize quantitative analyses and qualitative data from focus groups. The proposed research will entail a follow-up of a student cohort (N=184) at the University of Maryland, College Park, whose critical thinking abilities were previously assessed in their first year of college. A reassessment of these abilities in their fourth year will provide data to accomplish two goals: 1) an examination of the relationship between critical thinking and major now that they have moved beyond the more general introductory coursework, and 2) a longitudinal analysis of how their performance changed after several years of education. In addition, a sub-sample of students will be targeted for focus groups based on student characteristics such as socio-economic status, race/ethnicity, or commuter status. This strategy will allow investigators to gain student input about the broader outcomes they perceive, the factors they believe contribute to the outcomes, and reflective feedback on the assessment. This information will be used to guide further research on the student-level variables relevant to broad measures of student success.

The creation and testing of this test battery is the next critical step in developing a reliable, valid, generalizable measurement approach to the assessment of broad learning outcomes. The innovative aspects of the proposed study lie in its effort to address a fundamental question within higher education in particular and educational research more generally: How can we best measure and assess the broader analytical outcomes of an undergraduate education? Specific institutional learning contexts, such as major, as well as the contextual basis of the curricula (e.g., extracurricular involvement or commuter status) and student characteristics (e.g. race, SES) are identified as

patterns to observe rather than assume. Furthermore, critical thinking is investigated through multiple indicators and with varying strategies for observation, even within the same stimulus, and students are not expected to be best able to demonstrate their thinking and reasoning abilities with a single observational mode. These assumptions, which aim at greater specificity in the assessment design and analysis, will ultimately yield research that is able to make better and more accurate generalizations.

As a result of the proposed research, researchers and assessment specialists would have guidance for considering and accounting for student demographic and educational characteristics in the creation and evaluation of assessment measures. In addition, this work may provide insight into disciplinary approaches best promoting critical thinking outcomes in undergraduates. Further, focus group data will offer information to guide revision and/or expansion of this and similar assessments. In particular, student opinions will be sought about the relevance of the assessment for various student populations and about other broad outcomes that may be unique to specific student groups.

Once completed, the instrument will have a broad impact because it will be able to identify groups with greater and lesser differences in their critical thinking abilities—information that could be used for implementing curricular reform or instituting educational support programs. These data might also serve to address external accountability requirements originating from federal and state governments or accrediting organizations. In particular, improved awareness about the ways to achieve broader learning objectives can lead to educational policies that improve undergraduate learning experiences. Such improvements would yield benefits to undergraduate students and beyond, since thinking and reasoning abilities are widely recognized to be valuable for long-term educational, employment, and citizenship objectives.

Issues of instrument creation, construct measurement, and the factors associated with critical thinking and its development will be of interest to a wide variety of audiences. These include assessment specialists, higher education scholars, accreditation professionals, and those involved in the national higher education accountability dialogue. In addition, constituent groups of the Association of American Universities (presidents, provosts, and deans) have familiarity with the previous pilot studies and would appreciate a follow-up with more in-depth findings from this investigation. These individuals are empowered to implement on their own campuses assessment strategies that will grow out of this investigation, thereby creating the opportunity for more and broader institutional impacts from this study.

TABLE OF CONTENTS

A Longitudinal Study of Student Success: The Relation Between Academic Major, Student Demographics and Broad Student Outcomes

a. Statement of Problem

Part 1: Introduction and Overview

To what extent does an undergraduate education foster students' abilities to think critically and to reason? Similar to specific, curricular-centered learning outcomes or long-term earnings and employment prospects, student success as measured by the development of thinking and reasoning skills is among the most important goals of universities (U.S. Department of Education, National Center for Education Statistics [NCES], 2000a; 2000b). These abilities have lasting intellectual benefit because they are broadly applicable to students' longer-term educational and employment objectives (NCES, 1995). Further, a strong democracy requires intellectual give-and-take, communication, collaboration, deliberation, and a broadening of self-interest (Tocqueville, 1969; Putnam, 1995; Coleman, 1988), all of which hinge on the ability to reason and make critical judgments.

While the goal of critical thinking is common to most of higher education in the United States, institutions possess few tools and virtually no data to assess the extent to which graduating students demonstrate advancement in this area. The primary hurdle in collecting such information lies in the **definition** and **measurement** of the concept. Many scholars have defined critical thinking (Cottrell, 1999; Paul, 1990; Ennis, 1987; McPeck, 1981), but only a few have offered instruments or assessment items operationalizing their definition with student evaluation in mind. A major difficulty lies in the debate about the "disciplinarity" of critical thinking—that is, the extent to which critical thinking is a discipline-specific cognitive outcome, requiring measurement as such. This study addresses the assessment of a university's unique contribution to the development of broad student outcomes such as thinking and reasoning abilities by testing alternative measures both rooted within and spanning across academic disciplines. It may be the case for example, that students in the humanities have processes for thinking critically that best lend themselves to working on problems in those disciplines. It would follow that attempts to assess critical thinking outcomes for these students with assessments lacking humanities content could result in a good deal of measurement error.

While a few assessment tools measure critical thinking and reasoning in college populations, none have been routinely implemented throughout higher education (U.S. Department of Education, NCES, 2000a). Most target critical thinking with items only generally tied to curricular content in an effort to insure their relevance to a wide population (see for example, Watson-Glaser Critical Thinking Appraisal [WGCTA]), California Critical Thinking

Skills Test (CCTST), ACT Collegiate Assessment of Academic Proficiency (CAAP). An underlying assumption appears to be that broad "transferable" outcomes should be assessed with broad, widely-applicable measures. We posit that thinking and reasoning skills are developed through numerous educational experiences, some of which are discipline-specific (McPeck, 1992). We also assert that assessment cannot ignore contrasting disciplinary approaches, because students will draw on their curricular training to demonstrate critical thinking abilities. An instrument that *only* includes items generalizable to all undergraduates may not provide the best forum for students to deeply engage their thinking and reasoning abilities, nor may this be the best opportunity for observing and measuring them. Therefore, the items we have developed aim to measure a broad scope of student achievement without ignoring important student level variables such as academic major. We are now prepared to examine the items' relation to disciplinary training and student characteristics, the tasks to which this proposal is directed.

An important step in the creation of the assessment is to determine how performance on discipline-specific outcome measures differs from that on items which cross disciplinary boundaries. Therefore, one objective will be to ground critical thinking questions within disciplinary groupings (e.g., social science, engineering, the arts), and to match them with students' major field concentrations. Thus, this assessment does not rest on student familiarity with broad subject areas. Instead, it allows students to work both within their discipline and outside its boundaries. In so doing, this study will test whether critical thinking is triggered and best demonstrated when students can take full advantage of their disciplinary training with items grounded in their area of study.

A second objective of this study is to investigate the role of individual level characteristics in the achievement of critical thinking and other broad student outcomes. What factors are related to critical thinking outcomes (e.g. racial/ethnic status, socio-economic status, curricular/extracurricular activities, commuter status, etc.) What other broad educational outcomes are most readily identified by students (e.g. diversity awareness, open-mindedness, creativity, confidence)? Can preliminary patterns be identified between other outcomes and student level variables?

The proposed research is part of an on-going effort at the University of Maryland's Center for Assessment of Higher Education (CAHE) whose mission is to develop and foster collaborative, interdisciplinary research that focuses on assessment of the quality of education and research in higher education settings. The proposed study is part of an on-going examination of empirical measures of undergraduate student outcomes and will build on two earlier pilots at Columbia University and the University of Maryland, College Park. Supported by the Association of American Universities (AAU)—a group of 62 leading research universities in North America, Maryland's CAHE

was identified as the best location for continuing its investigation of student outcomes, which consisted of five years of intensive research. CAHE continues to work in collaboration with AAU staff and constituent members.

In these prior pilots, instruments with analytic essay questions and a battery of multiple choice items were devised and tested. Two types of items were developed: those that were generalizable **across** disciplines and those more firmly rooted **within** specific disciplines. Analyses of data from these projects found that **major field** of study significantly predicted performance on the assessment. In particular, students with math and engineering majors performed better on two essay questions with material that included a graph of data or experimental methods of hypothesis testing, even after controlling for SAT scores and GPA (Association of American Universities [AAU], 2004). The questions were not designed for students in particular majors, but rather attempted to present content expected to be familiar to students about to graduate from college. The better performance by students in certain majors provided the initial empirical evidence that major field influences performance on assessments of broad outcomes. It might be the case that training in science and engineering disciplines predisposes students to perform better on certain types of so-called "generic" critical thinking items. Or conversely, items with strong science- and math-based content may be more effective in stimulating critical thinking abilities in these students.

The proposed research will entail a **follow-up** of a freshman student cohort (N=184) at the University of Maryland whose critical thinking abilities were assessed in the pilot study noted above (AAU, 2004). A reassessment of these abilities in their senior year will provide data to accomplish two goals: 1) an examination of the relationship between critical thinking and major now that they have moved beyond the more general introductory coursework, and 2) a longitudinal analysis of how their performance changed after several years of education. To date, few studies have collected longitudinal data on critical thinking and reasoning abilities (Pascarella, Pierson, Wolniak & Terenzini, 2004) and none have done so with the goal of examining the role of disciplines in the development of these outcomes.

Unfortunately, sample sizes in the prior research were not adequate to quantitatively investigate most individual level demographic or educational variables in relation to performance on the assessment. Thus, this study proposes a multi-method strategy which will utilize quantitative analyses and qualitative data from focus groups. Powell (1996) defined a focus group as "a group of individuals selected and assembled by researchers to discuss and comment upon from personal experience, the topic that is the subject of the research." (p. 499). Among the advantages of focus groups are the in-depth quality of the information acquired, the relative efficiency of the

method, the potential benefit to participants who gain perspective and insight by exploring the subject with similar others, and the flexibility to follow-up on interesting leads or unanticipated material (Gibbs, 1997).

A sub-sample of students will be targeted for focus groups based on student characteristics such as socio-economic status, race/ethnicity, or commuter status. While these qualitative data will not enable broad generalizations about the larger student population, they will provide information to guide further research on the student level variables relevant to broad measures of student success. This strategy will allow investigators to gain student input about the broader outcomes they perceive, the factors they believe contribute to the outcomes, and reflective feedback on the assessment. As Kreuger (1988) has noted, focus groups can be effectively used in preliminary or exploratory stages of a study because they are particularly helpful in generating hypotheses and developing concepts for future investigation. In this case, they will also provide greater depth and detail about individual experiences and student outcomes than could be obtained from a written survey instrument.

Part 2: Contextual Framework

While there is a lack of **instrumentation** to measure critical thinking in college student populations, **critical thinking definitions** by educational psychologists, philosophers, and assessment specialists are abundant (see Johnson (1992) for a theoretical review). The Delphi report (Facione, 1990), characterizes critical thinking as "the process of purposeful, self-regulatory judgment . . . [that] gives reasoned consideration to evidence, context, conceptualizations, methods, and criteria." Cottrell's (1999) elaboration of critical thinking is also useful. For Cottrell, critical thinking involves considering issues carefully, evaluating evidence to support viewpoints, pondering the conclusions that would follow from particular viewpoints, and assessing their rationality.

From a different perspective, a major effort by the U.S. Department of Education National Center Education Statistics in 1995 sought to identify the relevant skills needed for critical thinking from the point of view of employers, policymakers, and academics. The goal was to assist the educational researcher in standardizing a definition that could be operationalized. Six behavioral skills were identified as central to critical thinking: interpretation, analysis, evaluation, inference, presenting arguments, and reflection. These elements formed the basis of the definition of critical thinking used for development of the assessment.

As noted earlier, there is a scholarly debate about the extent to which critical thinking is discipline-based (U.S. Department of Education, NCES, 1995, presents a concise review). In a few recent studies, more advanced students demonstrated greater degrees of critical thinking when compared to entering first-year students (Benjamin & Chun,

2003; AAU, 2004). Is this difference the result of their total educational experience (and maturity) or in-depth learning in their major? Other research has found that cognitive outcomes such as critical thinking are differentially affected by such student level characteristics as parental education, extracurricular involvement, employment, and peer interaction (Pascarella, Pierson, Wolniak & Terenzini, 2004). Overall, this research points to the need for greater specificity when assessing broad outcomes, with regard to both the creation of assessments and the composition of student samples.

Most faculty do not directly test for critical thinking outcomes. Traditional course-based assessments generally focus on the mastery of narrowly-defined concepts and curricular material related to the course. While faculty may argue they assess critical thinking with term papers, lab reports, student projects or theses, these are course linked, subjective measurements. What is needed is a generalizable, valid, reliable measure of critical thinking that can empirically determine whether or not transference of critical thinking abilities occurs and under what circumstances. This study contributes to these needs by utilizing three types of measures to examine where and when critical thinking transference might occur and what factors (e.g. demographic and educational) might contribute to it.

The measurement of critical thinking has found its way into two very different large-scale investigations of college student outcomes: the Collegiate Learning Assessment (CLA) of the RAND Corporation's Council for Aid to Education, and the National Survey of Student Engagement (NSSE) (Benjamin & Chun, 2003; NSSE, 2001). With different objectives than the research proposed here, they are the most recent attempts to develop instrumentation for measuring broad college student outcomes, including thinking and reasoning. The CLA seeks to measure thinking, reasoning, and written communication through a variety of objective measurement items originally developed for the GRE, the bar exam, and the New Jersey Department of Education. As a set, these items are closely correlated with the SAT (Klein, Kuh, Chun, Hamilton, & Shavelson, 2003). The CLA has a stated goal of **comparing institutions** with regard to their students' abilities, so institutions (rather than students) become the unit of analysis and reporting. Studies using the CLA instrument have not used a random sampling method. Instead, participation has been sought from small samples of students enrolled in specific courses. As a result of this design, inferences about non-participating individuals and subgroups of the student body are unreliable.

The NSSE takes a different approach by examining student outcomes through the means of self report. Using an opinion survey, students are queried about their participation in campus-based activities and their own evaluation of intellectual changes that have occurred as a result of their college experiences (both academic and non-academic).

NSSE data can inform university administrators about the strengths and weaknesses in curriculum, instruction, and campus life based on student perceptions (Kuh, 2001). Since numerous institutions participate, administrators can also make comparisons with selected peers. However, the self-reported nature of the data may cause some to consider it a less credible measure of college outcomes as abstract as critical thinking (NCES, 2000a).

Both of these assessments are designed with a one-size-fits-all approach, without accounting for individual characteristics among the student body. In contrast to the CLA effort, the proposed study will attempt to understand critical thinking at the **individual level**. In contrast to NSSE, this study will **directly measure critical thinking** by asking students to solve cognitive tasks. Thus, the project is significantly different than others currently underway nationally and will offer the opportunity to demonstrate the role of individual characteristics, including academic discipline, in student outcomes assessment.

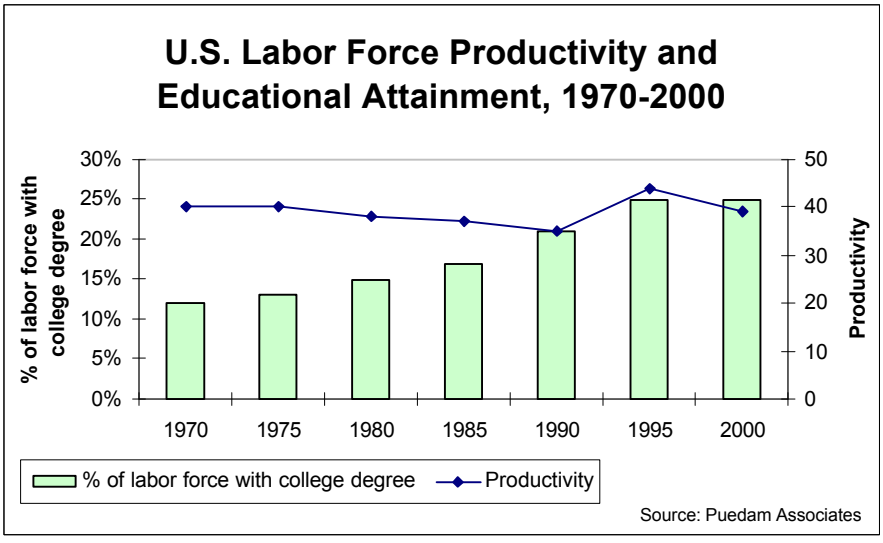
b. Proposal of Work

Part 1: Data Collection and Preparation

Three analytic essay items will be administered, one of which will be in the student's broad field area. Each essay prompt presents novel stimuli and requires students to make interpretations, derive meaning, draw conclusions, and make an argument. The essay questions were originally developed through consultation with Columbia University faculty, with modifications by faculty at UM, College Park. See an essay example below.

Write an essay providing an interpretation of the data in the figure below by answering the following questions.

a. What are the three most salient features of the data?
b. Name three additional pieces of information that would be helpful in forming your interpretation.
c. Imagine what the additional information might be. How would this knowledge alter your understanding of what is going on?
d. Given what you can observe, how are the data related, and what meaning can you infer?



In addition to the three analytic questions, the assessment includes portions of Sternberg's STAT (the Sternberg Triarchic Abilities Test), a multiple choice test designed to distinguish between academic-analytical, synthetic-creative, and practical-contextual intelligence. The test is comprised of a battery of standardized items that includes figural, quantitative, and verbal multiple-choice questions. The STAT places a greater emphasis upon learning than on memorization and is viewed as a useful benchmark for examining performance on the analytic questions.

The third component of the data collection will capture self-reported demographic and educational experience information. Additional data on students in the sample, including SAT scores, high school and college GPAs, majors or intended majors, and a full academic transcript, will also be made available through the campus' institutional research office. Analyses of these supplementary data will test for correlates of the critical thinking measures and will allow for a more complete interpretation of assessment results by providing extensive information about curricular experiences of participants. For example, engineering majors or students who had taken a course in reasoning or who had participated in an honors program could be disaggregated for separate analyses. Transcript data will also be used to validate that a student's declared major is indeed represented in their coursework.

The fourth and final component of the study will be the collection of focus group data. Because many important student level variables cannot often be examined quantitatively due to their low rates of occurrence in the student population, qualitative data will be collected through personal contact. Small focus groups will be employed for selected subgroups of students, such as those of a racial minority group, or who have a low socio-economic status, or who commute or are enrolled part-time, or who are of a non-traditional age. The specific composition of these groups will be determined upon evaluating the individual level variables of student participants.

For participation in the study, students will receive a cash incentive of $75 and will also be entered into a drawing to win a round-trip airline ticket within the continental United States. Cash incentives ranging from $20 to $40 were used during the pilot study at the University of Maryland in the fall of 2002. The examination of response rates for each incentive type found that $30 with a chance to win an airline ticket produced an average participation rate of 35%. These results are consistent with Salant and Dillman (1994) who found that the greater the incentive, the greater the return rate. Given the difficulty in conducting longitudinal studies and the importance of the UM sample, the investigators consider the cost well worth the investment. Indeed, if the $75/chance to win incentive

package would produce a 60 to 75% response rate, a great deal of confidence could be placed in the findings of the present study and extensive qualitative data could also be collected.

Recruitment procedures will also follow the method suggested by Salant and Dillman. Contacts will include email messages, letters, and phone calls. The utilization of multiple modes of communication provides the greatest likelihood that all of those in the sample are notified about the study more than once. Students will be recruited first through email, followed by a personalized letter, phone calls, and reminder email messages. The previous studies also demonstrated that personal contact with students was a factor positively contributing to response rates.

The 184 first-year students who participated in the pilot in 2002 will be recruited for this follow-up study. The institutional research office at the university will assist with providing contact information for the sample and supplemental student-level data for the analyses. The administration of the assessment will take place electronically. The services of the Massachusetts Institute of Technology's (MIT) Information System's Web Survey Service will be procured for programming of the instrument, electronic data collection, and sample management. Their work in facilitating these responsibilities for the previous pre-test and pilot at the University of Maryland was exemplary.

Project personnel will participate in the revision of scoring rubrics originally developed in the pilot study to standardized evaluations of essays across scorers. Since there is no consensus regarding the best scoring method for essays written for this purpose, three different scoring rubrics will be applied to each essay. **Holistic scoring** is a method that evaluates an essay in its entirety compared to specific pre-established criteria. It requires only one reading of each essay and is thus a fast, easy, and economical method, making it a very popular choice of scoring mechanism (Cooper, 1984). Holistic scoring is especially useful when an overall sense of quality is desired, when essays require an element of creativity, or when they do not have a definitive right or wrong answer (Mertler, 2001). It is also preferred for assessing complex and related skill sets that are difficult to compartmentalize (Westat, 2001).

Analytic scoring is most often contrasted with holistic scoring. This approach distinguishes between various components of an essay, assigning points or a grade for each component separately. While this method provides a stronger basis from which to provide diagnostic feedback to students, it also creates the impression that an essay is no more than the sum of its autonomous parts (Cooper, 1984; White, 1993). It can also be an impractical option as a separate reading might be required for scoring each component (Mertler, 2001). A third alternative, given much less attention in the literature, is **primary trait scoring**. This method assigns a score based on the most important trait

being assessed by the essay task. In this case the scorer would evaluate the essay in terms of the key construct, critical thinking, without consideration of other dimensions such as writing style or quality (Cooper, 1984).

Each type of scoring will be evaluated to determine the approach best suited for this study. The scoring rubrics will be pre-tested to determine their ease of use or need for revision. Each essay will then be assessed by three graders, who will include one of the principal investigators, one individual affiliated with the university writing program, and a third grader who is instructional staff at the university.

In addition to the three rubric analysis methods an emerging statistical approach called **latent semantic analysis** (LSA) will also be employed to score the essays and the results compared with those from the graders. LSA is a computerized method of analyzing large text by representing words and groups of words as points in a high dimensional space (Landauer, Foltz & Laham, 1998), and has been shown to estimate student knowledge as conveyed through essays (Wolfe et al., 1998; Rehder, Schreiner, Wolfe, Laham, Landauer & Kintsch, 1998). Its computerized scoring rubric produces contextual usage measures on word to word, word to passage and passage to passage bases. These measures can then be converted to holistic quality ratings in a number of ways as described by Landauer (1998). This statistical approach is particularly appealing in that it overcomes all subjective aspects of rater scoring and could be applied to large numbers of student essays at low cost.

The focus groups will be organized and conducted by the Principal Investigators and a graduate student funded under this proposal. The types of questions to be used can be described as "open ended probes." (Gibbs, 1993) Since the purpose is to elicit information without prescriptions, the questions will be worded so that they do not influence the participant's answer. Additionally, the questions should be clearly formulated with easier questions preceding more difficult ones. The sample questions below have been developed according to these criteria. The questions and session format will be piloted with volunteer students to determine the length, clarity of questions, and effectiveness in gaining information about broad student outcomes.

Sample Questions:

- Do you think college has changed how you think about things? In which ways?
- Has college influenced the way you think about people who are different from you? How so?
- Has college changed how you think/feel about yourself? How so?
- What about how you solve problems or deal with complicated issues? Has this changed about you since going to college? In which ways?

- What factor or experiences have contributed to these changes (a class, your major, individual faculty members, student peers, something else)?
- In talking with friends who've attended other colleges, do you find anything unique about how Maryland affected you?
- Compared to students attending other colleges, do you think anything unique about your background has contributed to a) what you brought with you to college? b) what college has provided?

Data from the pilot UM study was used to establish initial reliability and validity of the analytic essays and STAT items as well as the usefulness of the original scoring rubric. Internal consistency reliability using Cronbach's alpha was determined to be .72 for the STAT items. This is a good level of internal consistency for an instrument of this type. Inter-rater reliability was calculated to determine the consistency across essay scorers. Inter-rater reliability for the piloted study was in the range of .75, suggesting that trained raters can agree they identify the same features in a series of essays. As noted above, this study proposes to **extend the methods** used in scoring the essays and thus additional rubrics will be developed and piloted. Examinations of scoring rubrics will include correlating essay scores derived when essays are re-scored using different rubrics.

The pilot study examined validity of the measurement of the instrument battery using the following means. Since the analytic essays were originally reviewed by faculty members in a relevant discipline, a general sense of content validity was derived. These essays will be reviewed by additional content experts on UM's campus. Criterion validity was evaluated by using SAT scores and performance on the STAT to predict scores on the analytic questions. The correlation for STAT and verbal SAT was .51, for STAT and math SAT was .54, and for STAT and total SAT was .61. This moderate correlation suggests that the critical thinking battery taps into both general intelligence and unique critical thinking skills. Overall, the psychometric qualities of the items used in the pilot were found to be sufficiently robust that the investigators have confidence in their use for this study.

Part 2: Analyses

For the web-based assessment, the basic approach to data analysis draws on the techniques of structural equation modeling as well as conventional statistical tests such as analysis of covariance and *t* tests. Critical thinking is conceptualized as an unobserved, latent variable much like fundamental variables such as mass or acceleration that are familiar in the physical sciences. While the level of an individual's critical thinking skills cannot be directly observed, we have indirect measures, or indicators, that we believe reflect this unobservable level. For example, the Sternberg Triarchic Abilities Test (STAT) is composed of tasks (items) that reflect expert opinion about critical thinking. Thus, a total score or sub-scores from the STAT can be viewed as indictors of the level of

critical thinking skill for an individual. Data from the three analytical essay equations will also be scored on dimensions that are believed to represent critical thinking.

Figure 1 provides a schematic for the postulated latent variable structure. Ovals represent critical thinking that is viewed as having both a generalized component (larger oval labeled "critical thinking") and specific components associated with subject matter areas (smaller ovals labeled "essay 1," "essay 2" and "essay 3"). The gamma coefficients ($?_{A1}$, etc.) can be numerically larger or smaller depending upon the strength of the relation between the generalized component and subject-matter components of critical thinking. The first area of investigation will address the longitudinal component of the study and evaluate how these coefficients vary from the first to the second assessment period. As students progress from their first to senior year, their knowledge will begin to concentrate in their major area of study. We hypothesize that greater knowledge within a subject matter area results in greater ability to utilize critical thinking skills in that area. Support for this hypothesis would be found if the gamma coefficients at the second period are larger than at the first period, which will be assessed by a significance test.

A second emphasis will be on the relative contributions of general and subject-matter-specific components of critical thinking ability. Students will be grouped by subject-matter specializations (i.e., academic majors) and these contributions will be evaluated for generalizability. We hypothesize that when students receive an analytic question that is congruent with their general area of study, they will demonstrate greater subject specific critical thinking, and will also exhibit greater general critical thinking skills because their "transferable" reasoning will be stimulated by the familiar content. If content does not matter for general critical thinking abilities, we would observe no difference in the gamma coefficients for the essay questions despite the contrasting focus of the items. Overall, these analyses will offer empirical evidence about the generalizability of critical thinking skills.

To assist in the analysis of the qualitative data collection, the focus groups will be tape-recorded. The tapes will be transcribed and used for a content analysis of our key questions. A final report will be written that outlines the results and the implications for the development of the assessment instrument.

Pending the necessary funding, the proposed study will be undertaken over a twelve month time period, from June 2005 through June 2006 (see Table 1 for Plan of Work).

c. Dissemination Plan

The findings from the study will be disseminated in numerous fields and through a variety of presentation and publication outlets. Within the field of higher education, scholarly outlets will include the American Educational

Research Association, the Association for the Study of Higher Education and the AIR/NPEC Conference on Student Success. In addition, the American Psychological Association also has members who would value the information provided by this study. Presentations at meetings of these associations will form the basis of scholarly manuscripts that will be prepared. In addition, the most promising findings will be disseminated to the presidents, provosts, and deans of other research universities through contact with constituent groups of the Association of American Universities (AAU).

d. Description of Policy Relevance

This work may provide insight into disciplinary approaches best promoting critical thinking outcomes in undergraduates. As a result of the proposed research, researchers and assessment specialists would have guidance for considering and accounting for student demographic and educational characteristics in the creation and evaluation of assessment measures. Further, focus group data will offer information to guide revision and/or expansion of this and similar assessments. In particular, student opinions will be sought about the relevance of the assessment for various student populations and about other broad outcomes that may be unique to specific student groups. Once completed, the instrument will have a broad impact because it will be able to identify groups with greater and lesser differences in their critical thinking abilities—information that could be used for implementing curricular reform or instituting educational support programs. Under certain circumstances, these data might also serve to address external accountability requirements originating from federal and state governments or accrediting organizations such as ABET. In particular, improved awareness about the ways to achieve broader learning objectives can lead to educational policies and reforms that improve undergraduate learning experiences. Such improvements would yield benefits to undergraduate students and beyond, since thinking and reasoning abilities are widely recognized to be valuable for long-term educational, employment, and citizenship objectives.

e. Discussion of innovative aspects of project

The innovative aspects of the proposed study lie in its effort to address a fundamental question within higher education in particular and educational research more generally: How can we best measure and assess the broader analytical outcomes of an undergraduate education? In this regard, the proposed research will advance knowledge about student outcomes assessment, and will especially address the role of disciplines in developing or promoting critical thinking and reasoning abilities. In order to make progress in understanding the dynamics of the relationship between undergraduate educational experiences and broad measures of student success, we do not assume these

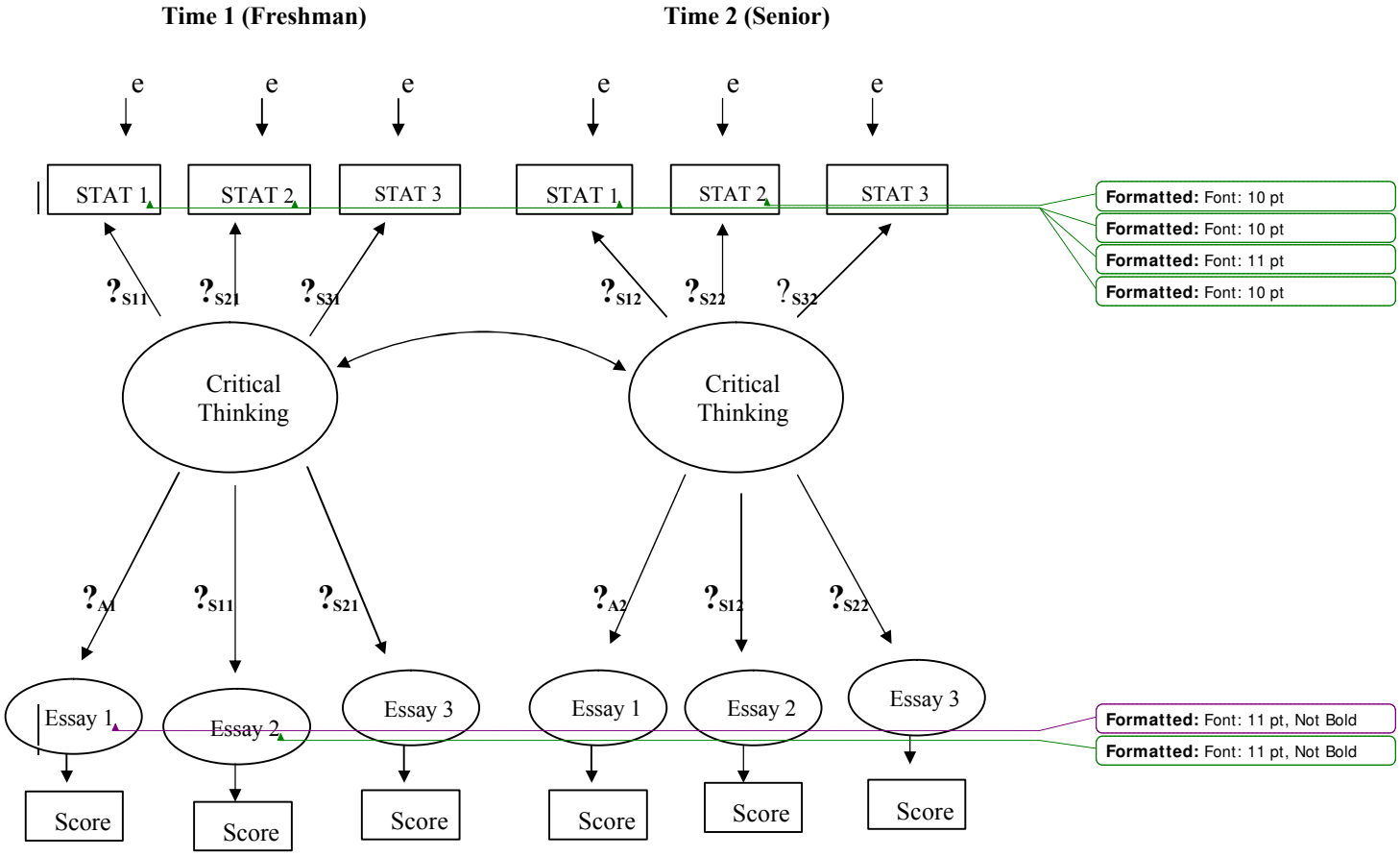
experiences to operate similarly for all students, or that a college education produces similar benefits for all students, even at a single institution. Thus, specific institutional learning contexts, such as major, as well as the contextual basis of the curricula (e.g., extracurricular involvement or commuter status) and student characteristics (e.g., race, SES) are identified as patterns to observe rather than assume. Furthermore, critical thinking is investigated through multiple indicators and with varying strategies for observation even within the same stimulus, and students are not expected to be best able to demonstrate their thinking and reasoning abilities with a single observational mode. These assumptions, which aim at greater specificity in the assessment design and analysis, will ultimately yield research able to make better and more accurate generalizations.

The creation and testing of this test battery will provide the next critical step in developing a reliable, valid, generalizable measurement approach to the assessment of broad learning outcomes. Additionally, the proposed study also benefits from a wide range of expertise from the co-PIs at the University of Maryland, College Park, including measurement and assessment specialists and psychometric testing experts. The PIs experience with the previous pilot studies, combined with the faculty and staff at the institution, yields a team well-equipped to accomplish the objectives of the proposed research.

f. Discussion of audience to whom the project will be important

Issues of instrument creation (validity, reliability), construct measurement (raters, rubrics, and LSM), as well as the factors associated with critical thinking and its development will be of interest to a wide variety of audiences, including assessment specialists, higher education scholars, accreditation professionals, and those involved in the national higher education accountability dialogue. In addition, constituent groups of the AAU (presidents, provosts, and deans) have familiarity with the previous pilot studies and would appreciate a follow-up with more in-depth findings from this proposed investigation. These individuals are empowered to implement on their own campuses assessment strategies that will grow out of this investigation, thereby creating the opportunity for more and broader institutional impacts from this study. The goal is to begin creating a network of faculty and administrators who become aware of the potential usefulness of this assessment battery and perhaps contribute to the dialogue and investigation of broad student outcomes, beginning with critical thinking. Interested scholars and assessment specialists at other institutions will be encouraged to use this material at their own institutions—especially those that may offer a very different learning environment such as HBCUs and liberal arts colleges.

Figure 1: Conceptual Model for Analysis



NOTES: STAT 1-3 are sub-scores from the Sternberg Triarchic Abilities Test. Essay 1-3 refer to the analytical essay questions; all students completed essay 1 while subsets of students completed essays 2-3. Thus, the analyses involve missing observations.

Table 1: Plan of Work

June – Aug 2005	*Primary Personnel*	*Sept – Dec 2006*	*Primary Personnel*	*Jan – April 2006*	*Primary Personnel*	*May – June 2006*	*Primary Personnel*
Develop Scoring Rubrics and Focus Group Protocol a. Consult with instrument and rubric reviewers b. Develop rubrics (3 forms) c. Pretest and revise rubrics d. Develop and test focus group discussion guide e. Revise guide	Kivlighan, Schmidt, Brooks	**Data Collection** a. Coordinate and test electronic coding of assessment b. Establish recruitment schedule and generate correspondence c. 10 day field period d. Monitor data collection and response rates e. Distribute incentives	Brooks, Grad RA	**Score Essay Data** a. Coordinate raters b. Apply LSA method to data c. Enter scores into data file for analysis	Brooks, Grad RA	**Research Dissemination** a. Prepare conference proposals b. Paper and manuscript preparation c. Consult with AAU and CAHE advisory board members	Kivlighan, Brooks, Schmidt
Rescore Prior Data a. Provide training to raters (3 methods) b. Test alternative LSA methods c. Generate scores d. Code into data file e. Analyze relationships among scoring methods	Kivlighan, Schmidt, Brooks,	**Conduct Focus Groups** a. Select and contact participants b. 10 day field period	Brooks, Schmidt, Grad RA	**Analyses** a. Run models using regression analysis & structural equation modeling b. Analyze focus group transcripts c. Prepare focus group report	Kivlighan, Brooks, Grad RA		
		Data Cleaning a. Establish completion criteria for case inclusion b. Examine outliers c. Examine missing data	Kivlighan, Schmidt, Brooks				

REFERENCES CITED

Association of American Universities, Assessing Quality of University Education and Research Project. (2004, March). *Critical thinking and analytical reasoning among undergraduate students: A pilot study at the University of Maryland, College Park.* Unpublished report.

Benjamin, R. & Chun, M. (2003, Summer). A new field of dreams: The Collegiate Learning Assessment Project. *Peer Review*, 26-29.

Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, *94*(Suppl.), S95-S120.

Cooper, P. L. (1984, May). *The assessment of writing ability: A review of research* (GRE Board Research Rep. No. 82-15R) [Electronic version]. Available from Graduate Records Examination Web site.

Cottrell, S. (1999). *The Study Skills Handbook.* MacMillan, London: 188-199.

Ennis, R. H. (1987). A taxonomy of critical thinking dispositions and abilities. In J. B. Baron & R. J. Sternberg (Eds.), *Teaching thinking skills: Theory and practice.* New York: Freeman.

Facione, P. A. (1990). *The Delphi report ~ Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction.* Millbrae, CA: The California Academic Press. (ERIC Document Reproduction Service No. ED315423)

Gibbs, A. (1997). Focus groups. *Social Research Update, 19*. Conducted for the Department of Social Medicine at Briston University.

Gibbs, G. (1993). *Learning in teams: A student guide.* Oxford, England: Oxford Brooks University Press.

Johnson, R. H. (1992). The problem of defining critical thinking. In S. Norris (Ed.), *The generalizability of critical thinking: Multiple perspectives of an educational ideal*. New York: Teachers College Press.

Klein, S., Kuh, G., Chun, M., Hamilton, L., & Shavelson, R. (2003, April). *The search for value-added assessing and validating selected higher education outcomes*. Paper presented at the 84th annual meeting of the American Educational Research Association, Chicago, IL.

Kreuger, S. (1988). *Focus group: A practical guide for applied research*. London: Sage Publishers.

Kuh, G. D. (2001). Assessing what really matters to student learning: Inside the National Survey of Student Engagement. *Change*, *33*(3), 10-17, 66.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes*, *25*(2&3), 259-284.

McPeck, J. E. (1992). *Thoughts on subject specificity.* In S. Norris (Ed.), *The generalizability of critical thinking: Multiple perspectives of an educational ideal.* New York: Teachers College Press.

McPeck, J. E. (1981). *Critical thinking and education.* New York: St. Martin's.

Mertler, C. A. (2001). Designing scoring rubrics for your classroom. *Practical Assessment, Research & Evaluation*, *7*(25). Retrieved October 8, 2003, from http://edresearch.org/pare/getvn.asp?v=7&n=25

National Survey of Student Engagement. (2001). *Improving the college experience: National benchmarks of effective educational practice* (NSSE 2001 Report). Bloomington, IN: Indiana University Center for Postsecondary Research and Planning.

Pascarella, E. T., Pierson, C. T., Wolniak, G. C., & Terenzini, P. T. (2004, May/June). First-generation college students: Additional evidence on college experiences and outcomes. *The Journal of Higher Education*, *75*(3), 249-284.

Paul, R. W. (1990). *Critical thinking: What every person needs to survive in a rapidly changing world*. Rohnert Park, CA: Center for Critical Thinking and Moral Critique, Sonoma State University.

Powell, R. A. & Single H.M. (1996). Focus groups. *International Journal of Quality in Health Care*, *8* (5), 499-504.

Putnam, R. D. (1995). Bowling alone [Electronic version]. *Journal of Democracy*, *6*(1), 65-78.

Rehder, B., Schreiner, M. E., Wolfe, M. B. W., Laham, D., Landauer, T. K., & Kintsch, W. (1998). Using latent semantic analysis to assess knowledge: Some technical considerations. *Discourse Processes*, *25*(2&3), 337-354.

Salant, P. & Dillman, D. A. (1994). *How to conduct your own survey*. Indianapolis, IN: Wiley.

Tocqueville, A. de. (1969). *Democracy in America* (J. P. Maier, Ed., & G. Lawrence, Trans.). Garden City, NY: Anchor Books.

U.S. Department of Education, National Center for Education Statistics. (2000a). *The NPEC sourcebook on assessment, volume 1: Definitions and assessment methods for critical thinking, problem solving, and writing* (Report No. NCES 2000--172). Prepared by T. D. Erwin for the Council of the National Postsecondary Education Cooperative Student Outcomes Pilot Working Group: Cognitive and Intellectual Development. Washington, DC: U.S. Government Printing Office.

U.S. Department of Education, National for Center Education Statistics. (2000b). *The NPEC sourcebook on assessment, volume 2: Selected institutions utilizing assessment results* (Report No. NCES 2000-172). Prepared by T. D. Erwin for the Council of the National Postsecondary Education Cooperative Student Outcomes Pilot Working Group: Cognitive and Intellectual Development. Washington, DC: U.S. Government Printing Office.

U.S. Department of Education, National Center Education Statistics. (1995). *National assessment of college student learning: Identifying college graduates' essential skills in writing, speech and listening, and critical thinking* (Report No. ISBN-0-16-048051-5). Washington, DC: U.S. Government Printing Office. (ERIC Document Reproduction Service No. ED383255)

White, E. M. (1993). Assessing higher-order thinking and communication skills in college graduates through writing. *JGE: The Journal of General Education*, *42*, 2.

Wolfe, M. B. W., Schreiner, M. E., Rehder, B., Laham, D., Foltz, P. W., Kintsch, W. et. al. (1998). Learning from text: Matching readers and texts by latent semantic analysis. *Discourse Processes*, *25*(2&3), 309-336.

BIOGRAPHICAL SKETCH

Rachelle L. Brooks, Ph.D.
Associate Director
Center for Assessment of Higher Education
University of Maryland, College Park

Dr. Rachelle Brooks began developing her skills in quantitative methodology as a political science major at the College of Wooster, Wooster, Ohio. While there, she also developed strong analytical skills and pursued advanced coursework in numerous social science disciplines. This more general interest in social science research, as it is broadly conceived, continued even as she pursued her Ph.D. at Rutgers University, New Brunswick, New Jersey. In addition to conducting research in political science pursuant to her doctoral degree, she also worked as a researcher for several different projects, addressing issues as diverse as healthcare, workforce development, and undergraduate leadership training. While in graduate school she also sought additional formal methodological training at the ICPSR Summer Program in Quantitative Methods for Social Research. This experience exposed her to quantitative methods techniques not readily available in her department, including advanced linear regression techniques, maximum likelihood estimation, scaling and dimensional analysis, and simultaneous equation modeling.

At Rutgers, Dr. Brooks' dissertation research required the analysis of a large national database that included additional follow-up information on a sub-sample of respondents. By linking the screener and follow-up data she was able to estimate a series of two-stage least squares regression models to complete her analysis. Her work at the Center for Public Interest Polling (CPIP) at Rutgers also gave her experience analyzing qualitative data, including focus groups, interview transcripts, and journal entries.

Additional experience working with large databases was gained while she was employed as research manager and consultant for Consumer Health Sciences in Princeton, New Jersey. Dr. Brooks managed a longitudinal database that tracked caregiver and patient experiences at six-month intervals. While responsible for many aspects of the research, from survey development and administration to data collection and analysis, she found most rewarding the analysis of the data and the development of manuscripts based on her results.

After completing her doctoral degree in Political Science, Dr. Brooks joined the Association of American Universities to develop and implement a 5-year pilot project to develop improved measures of the quality of university education and research that will aid in more accurate and informative institutional descriptions and assessments. As the Director of Research and former Project Manager, Dr. Brooks designed and conducted faculty opinion surveys as well as student outcomes assessments, and also undertook a secondary analysis of a national

reputational assessment of doctoral programs. This was all with the aim of better understanding the meaning and measurement of quality in postsecondary institutions. In addition, Dr. Brooks conducted eight in-depth campus field visits to better understand how data is collected, accessed, and used by the administration of postsecondary institutions for the purposes of planning and policy-making. Dr. Brooks also works closely with the Association of American Universities Data Exchange, and has carefully monitored their impressive progress in developing a data warehouse to allow research universities to share institutional level data.

Her work at the AAU was identified by the University of Maryland, College Park, as meriting continuation through the establishment of a new research center—the Center for Assessment of Higher Education—where Dr. Brooks is currently the Associate Director. In partnership with senior faculty she is establishing the Center's research priorities. See below for excerpts of her C.V.

EDUCATION

Ph.D., Political Science (Major: Women and Politics), 2000; Rutgers University, New Brunswick, New Jersey

Dissertation: "The Effects of Paid Employment on Women's Political Activity: An Analysis of Structural and Contextual Workplace Factors," supervised by Susan J. Carroll and Jane Junn

- Graduate Certificate in Women's Studies
- ICPSR Summer Program in Quantitative Methods of Social Research; University of Michigan, Ann Arbor

M.A., Political Science, 1996; Rutgers University, New Brunswick, New Jersey

B.A. 1990; College of Wooster, Wooster, Ohio; (Major: Political Science; Minor: Women's Studies)

Thesis: "The Political Party Strategies of the Contemporary Feminist Movement," supervised by Karen Beckwith

RESEARCH ACTIVITIES

Associate Director, *Center for Assessment of Higher Education (CAHE)*
University of Maryland, College Park, MD, 2004-present.

- Assisted dean and center directors to establish priorities for a newly established research center.
- Organized advisory board to advise and oversee work of the center
- Prepared grant proposals to federal agencies and private foundations to support the research of the center.

Director of Research, *Assessing Quality of University Education and Research (AQUER) project*, *Association of American Universities*
Washington, D.C., 2002-2004.

- Designed and implemented pilot studies to improve upon the measurement of the quality of university research and education programs, using online research methods for faculty opinion surveys and student outcomes assessments.
- Responsible for all external communications of the project's research, including presentations at higher education research and policy conferences, and internal research reports.
- Supervisor of project research assistants and administrative staff.

Project Manager, *Assessing Quality of University Education and Research (AQUER) project*, *Association of American Universities*
Washington, D.C., 2000-2002.

- Assisted in early planning of project goals and constitution of oversight committees for 5-year pilot project.
- Organized 2-day annual conferences of advisory committees comprised of university presidents, provosts, planning officers, and distinguished faculty, to devise a research agenda and report on research studies.
- Responsible for research staff hiring and budget management.
- Wrote or edited and disseminated all project correspondence.

Research Manager, *Consumer Health Sciences,* 1997-1998
Consultant, *Consumer Health Sciences,* 1998-2000
Princeton, NJ

- Managed a longitudinal database, conducted multivariate statistical analyses, and authored research reports and peer-review journal articles in the health science field.
- Coordinated all aspects of the research process, including questionnaire development, database creation, statistical analyses, and development of manuscripts, reports and presentations.
- Responsible for training and supervision of research assistants.

Research Associate, *Heldrich Center for Workforce Development*
Rutgers University, New Brunswick, NJ, 1998-2000.

- Researched and authored several components of literature review for US Department of Labor's 5-Year Research Plan Draft, including all DOL published and contracted research for 10-year period.
- Solicited suggestions from various federal departments regarding research priorities to include in plan.
- Assisted at meetings held in Washington, DC, for national experts in workforce development, by synthesizing their comments and suggestions for draft of research plan.

Research Assistant, *Center for Public Interest Polling, Eagleton Institute of Politics*
Rutgers University, New Brunswick, NJ, 1996-1997.

- Analyzed quantitative and qualitative opinion data.
- Authored reports in the form of statistical summaries and qualitative analyses for journalistic and research purposes.

PUBLICATIONS/PRESENTATIONS

Higher Education
Brooks, R. "Measuring University Quality," *Review of Higher Education,* forthcoming.

Junn, J. and Brooks, R. "How Well Do Reputational Assessments Approximate the Quality of Ph.D. Programs," in progress, research funded by the Spencer Foundation.

Brooks, R., Nerad, M., Morrison, E., and Cerny, J. "Ph.D. Employment Outcomes as a Measure of Program Quality," in progress.

"Faculty Scholarship in the Humanities: An Examination of Research Processes and Products," manuscript in progress, prepared for the American Academy of Arts & Sciences Humanities Indicators project.

"Changes in Faculty Scholarly Activity Over Time: Faculty Orientations, Research Processes and Research Products," manuscript in progress, research funded by the American Educational Research Association.

"Measuring Faculty Scholarly Activity," roundtable discussion, American Association for Higher Education Learning to Change Conference, Washington, DC, March 14-17, 2003.

"Assessing Quality of University Education and Research Project," with Jenniffer Manning, prepared for the 14th International Conference on Assessing Quality in Higher Education, Vienna, Austria, July 22-24, 2002.

"Undergraduate Student Outcomes as Value-Added: How Might Critical Thinking and Analytic Reasoning Skills Be Measured Better?" prepared for the American Association for Higher Education Assessment Conference, Boston, MA, June 19-23, 2002.

"How Well Can Reputations Be Measured? Analysis of 1992-93 NRC Data," with Jane Junn, delivered to the Association for Institutional Research, 42nd Annual Forum, Toronto, Canada, June 2-5, 2002.

"How Can Scholarly Productivity Be Measured Better? Results from an On-line Faculty Opinion Survey," delivered to the Association for Institutional Research, 42nd Annual Forum, Toronto, Canada, June 2-5, 2002.

Political Science/Women's Studies

"An Analysis of the Violence Against Women Act," in *Feminists Negotiate the State: The Politics of Domestic Violence,* Cynthia Daniels, editor. University Press of America, 1997.

"The Effects of Paid Employment on Group Consciousness and Political Participation: A Broadening of the Role of Work in Political Life," poster session prepared for the Annual Meeting of the American Political Science Association, Washington, D.C., August 1997.

"The Effects of Gender-Segregated Employment on Women's Political Participation," delivered to the Institute for Research on Women's Celebration of Our Work Conference, New Brunswick, NJ, May 1996.

"Campaigning for Resources: Political Parties Respond to Candidate-Centered Elections," delivered to the Annual Meeting of the Midwest Political Science Association, Chicago, IL, April 1996, and to the Political Science Graduate Student Association Conference at Rutgers University, March 1996.

"The Influence of Paid Employment on Women's Political Participation," delivered to the Annual Meeting of the Midwest Political Science Association, Chicago, IL, April 1995, and to the Political Science Graduate Student Association Conference at Rutgers University, March 1995.

"The Political Party Strategies of the Contemporary Feminist Movement," delivered to the National Conference on Undergraduate Research, Schenectady, NY, April 1990.

Health Sciences

"An Economic Analysis of Donepezil in the Treatment of Alzheimer's Disease," with Gary Small, MD, and Jane A. Donohue, PhD, *Clinical Therapeutics,* July/August, 1998.

"The Relationship Between Donepezil and Behavioral Disturbances in Patients with Alzheimer's Disease," with Jeffrey L. Cummings, MD, and Jane A. Donohue, PhD, *American Journal of Geriatric Psychiatry,* Vol. 8, May 2000.

"The Impact of Donepezil on Caregiving Burden for Patients with Alzheimer's Disease," with Howard M. Filit, MD, and Elane M. Gutterman, PhD, *International Psychogeriatrics,* Vol. 12, No. 3, September 2000.

INVITED PANELS AND BOARD MEMBERSHIPS

Teagle Foundation, “Listening Conference on Value-Added in Undergraduate Education,” invited participant, Little Switzerland, NC, September 3-6, 2004.

American Council on Education, "Roundtable for International Education Researchers," invited participant, Baltimore, MD, May 28, 2004.

Center for Innovation and Research on Graduate Education, University of Washington, Seattle; Survey Advisory Board Member, 2003-present.

American Academy of Arts & Sciences, "Humanities Indicators Research Issues Workshop," invited panelist, Cambridge, MA, January 9-10, 2003.

Council on Research Policy and Graduate Education, National Association of State Universities and Land-Grant Colleges Annual Meeting, "Accountability Measures, Ranking Criteria, and the Upcoming NRC Study," invited panelist, Washington, DC, November 12, 2001.

Northeast Association of Graduate Schools Annual Meeting, "Rankings and Assessment in Graduate Education," invited panelist, Annapolis, MD, March 31, 2001.

BIOGRAPHICAL SKETCH

Dennis M. Kivlighan, Jr., Ph.D.
Professor and Chair
Department of Counseling and Personnel Services
College of Education
University of Maryland, College Park

Dr. Dennis Kivlighan's research experience spans 20 plus years. During this 20-plus year research career, he is the fifth most published author in his specialty's leading empirical journal, *The Journal of Counseling Psychology*. Dr. Kivlighan is one of the leading research methodologists in the counseling arena. He has introduced several statistical and methodological innovations to the counseling field. In addition to statistics and methodology classes during his Ph.D. preparation, Dr. Kivlighan received postdoctoral training in Structural Equation Modeling. With his co-authors, Dr. Kivlighan published two editions of *Research Design in Counseling*, the best selling research design text among counseling programs nationally and internationally. At the University of Missouri, Dr. Kivlighan developed and taught a number of research and statistics courses at the doctoral level. These courses include: Research Design and Methodology, Advanced Regression Analysis, Hierarchical Linear and Nonlinear Modeling, and Advanced Multivariate Statistics.

In addition to his work in research design and methodology, Dr. Kivlighan's content expertise is in group counseling and group dynamics. He is the current editor of *Group Dynamics: Theory, Research and Practice*, the official journal of Division 49: Group Psychology and Group Psychotherapy of the American Psychological Association. He is very knowledgeable in the area of empirically examining important group process variables. For this specific research, Dr. Kivlighan brings extensive experience conducting focus groups and in analyzing data from focus groups.

A recent focus of Dr. Kivlighan's scholarship is the examination of expertise. In particular, he has used sophisticated Quantitative and qualitative methodology to examine the differences in how experts and novices structure knowledge. Along with others, Dr. Kivlighan's research suggests that knowledge structure, rather than knowledge content, is the critical element defining the superior performances on experts when compared to novices. As the development of critical thinking skills is an instance of the acquisition of knowledge structures, Dr. Kivlighan's research on knowledge structures and expertise will contribute significantly to this proposed project.

Dr. Kivlighan is currently the Co-Director for the Center for Assessment in Higher Education. He is involved in the development of measures to assess the impact of undergraduate education. His research also

involves examining the relationship between the assessment of the quality of teaching and the learning outcomes in classes. Selected elements of Dr. Kivlighan's C.V. are included on the following pages.

EDUCATION

Ph.D.	Virginia Commonwealth University, August 1982 Counseling Psychology (APA Accredited) Specialty Area: Group Counseling and Psychotherapy
M.S.	Virginia Commonwealth University, December 1980 Counseling Psychology
B.S.	College of William and Mary, January 1975 Psychology

Professional Experience

2001-Present	Professor and Chair, Department of Counseling and Personnel Services, University of Maryland, College Park.
1999-2001	Professor and Chair, Department of Educational and Counseling Psychology, University of Missouri, Columbia.
1986-1999	Professor, Associate Professor, Director of Training and Director of Graduate Studies, and Assistant Professor, Department of Educational & Counseling Psychology, University of Missouri, Columbia

PROFESSIONAL ACTIVITIES

- Editor: Group Dynamics: Theory, Research, and Practice
- Editorial Boards: Journal of Counseling Psychology, Psychotherapy Research
- Co-Chair: Division 17 Hospitality Suite APA Convention
- Membership Chair: Division 17 of APA
- Board of Directors, Council of Counseling Psychology Training Programs
- Consultant, Missouri Department of Family Services

HONORS

- Fellow, (Counseling Psychology) American Psychological Association
- Fellow (Group Psychology/Group Psychotherapy) American Psychological Association
- Research Publication Award, 1996, Association for Specialists in Group Work
- High Flyer Teaching Award, 1994-1998, College of Education, University of Missouri-Columbia.
- Honorable Mention, Research Publication Award, 1994, Association for Specialists in Group Work.
- Honorable Mention, Research Publication Award, 1994, Association for Specialists in Group Work.
- Research Publication Award, 1991, Association for Specialists in Group Work.
- Gold Chalk Award, for Excellence in Graduate Teaching, 1989, Graduate and Professional Council, University of Missouri, Columbia.
- Honorable Mention, Research Publication Award, 1984, Association for Specialists in Group Work

BOOKS AND CHAPTERS

Kivlighan, D. M., Jr., Holmes, S. E. (2004). The importance of therapeutic factors: A typology of therapeutic factors studies. In J. L. DeLucia-Waack, D. A. Gerrity, C. R. Kalonder, & M. T. Riva (Eds.) *Handbook of group counseling and psychotherapy* (pp. 23-36). Sage Publications, Inc., Thousand Oaks, CA.

Kivlighan, D. M., Jr. (2002). Transference, interpretation, and insight: A research-practice model. In G. S. Tryon (Ed.), *Counseling based on process research: Applying what we know* (pp. 166? 196). Boston: Allyn and Bacon.

Kivlighan, D. M., Jr., Coleman, M. N., & Anderson, D. C. (2000). Process, outcome and methodology in group counseling research. In S. D. Brown & R. W. Lent (Eds.). *Handbook of counseling psychology*, (3rd ed.) New York: Wiley, pp. 767-796.

Kivlighan, D. M., Jr., (2000). *Counseling process and outcome.* In A. E. Kazdin (Ed.). Encyclopedia of psychology. Washington, DC: American Psychological Association

Heppner, P.P., Kivlighan, D. M., Jr., & Wampold, B. E. (1999). *Research design in counseling* 2nd Ed. Pacific Grove, CA: Brooks/Cole.

Heppner, P.P., Kivlighan, D. M., Jr., & Wampold, B. E. (1992). *Research design in counseling*. Pacific Grove, CA: Brooks/Cole.

Kivlighan, D. M., Jr. (1990). Counseling theorists: Reflections in a mirror. In Heppner, P.P. *Pioneers in counseling and human development*. Alexandria, VA: American Association of Counseling and Development.

RECENT ARTICLES

Mallinckrodt, B., Porter, M. J. & Kivlighan, D. M., Jr. (in press). Client attachment to Therapist, Depth of In-Session Exploration, and Object Relations in Brief Psychotherapy. *Psychotherapy Research*

Kivlighan, D. M., Jr., & Luiza, J. W. (in press). Examining the credibility gap and the mum effect: Rex Stockton's contributions to research on feedback in counseling groups. *Journal for Specialist in Group Work.*

Kivlighan, D. M., Jr., & Kivlighan, M. C. (2004). Counselor intention use in individual and group treatment. *Journal of Counseling Psychology, 51*, 347-353.

Duan, C., & Kivlighan, D. M., Jr. (2002). Relationships among therapist presession mood, therapist empathy and session outcome. *Psychotherapy Research, 12*, 23-37.

Kivlighan, D. M., Jr., & Tarrant, J. M. (2001). Does group climate mediate the group leadership-group member outcome relationship: A test of Yalom's hypotheses about leadership priorities. *Group Dynamics: Theory, Research, and Practice, 5*, 220-234.

Holmes, S. E., & Kivlighan, D. M., Jr. (2000). Therapeutic factors in group and individual treatment. *Journal of Counseling Psychology, 47*, 478-484.

Kivlighan, D. M., Jr., & Shaughnessy, P. (2000). Patterns of Working Alliance Development: A Typology of Working Alliance Ratings. *Journal of Counseling Psychology, 47*, 362-371.

Kivlighan, D. M., Jr., Multon, K. D., & Patton, M. J. (2000). Insight and Symptom Reduction in Time-Limited Psychoanalytic Counseling. *Journal of Counseling Psychology, 47*, 50-58.

Kivlighan, D. M., Jr., & Arthur, E. G. (2000). Convergence of counselor-client recall of session critical incidents. *Journal of Counseling Psychology, 47*, 78-84.

Mayfield, W. A., Kardash, C. M., & Kivlighan, D. M., Jr. (1999). Differences in case conceptualization between experienced and novice counselors. *Journal of Counseling Psychology, 46*, 504-514.

Kivlighan, D. M., Jr., & Coleman, M. N. (1999). Values, exchange relationships, group composition and leader-member differences: A potpourri of reactions to Dose (1999). *Group Dynamics: Theory, Research, and Practice, 3*, 33-39.

Heppner, M. J., Neville, H. A., Smith, K., Kivlighan, D. M., Jr., & Gershuny, B. S. (1999). Examining immediate and long-term efficacy of rape prevention programming with black and white fraternity men. *Journal of Counseling Psychology, 46*, 16-26.

Brossart, D. F., Willson, V. L., Patton, M. J., Kivlighan, D. M., Jr., & Multon, K. D. (1998). A time series model of the working alliance: A key process in short-term psychoanalytic counseling. *Psychotherapy, 35*, 197-205.

BIOGRAPHICAL SKETCH

Janet A. Schmidt, Ph.D.
Director of Interdisciplinary Research
College of Education
University of Maryland, College Park

Dr. Janet Schmidt has been involved with research concerning college students since her career began at the University of Maryland. She coordinated a large, longitudinal study of college students from freshmen to seniors (Maryland Longitudinal Study-MLS) as well as the 10 year follow-up of this same group of students. Dr. Schmidt has extensive experience in the design of surveys, survey analysis and data interpretation. In 1995, she became the University of Maryland's Director of Institutional Research. In this capacity, she was involved with institutional data and national state reporting, and continued collaboration with Student Affairs, initiating joint research efforts. In 1988, Dr. Schmidt was appointed Director of Engineering Student Research at the Clark School of Engineering. There she played a key role in the college's original ABET EC 2000 accreditation which stressed assessment of student learning outcomes as well as using data for the improvement of student learning. She also has considerable experience conducting focus groups with undergraduate and graduate students for various NSF projects. Currently, Dr. Schmidt is responsible for initiating interdisciplinary research projects within the College of Education and across the University, as well as for faculty development activities related to research productivity. Her C.V. is included on the following pages.

JANET A. SCHMIDT

3611 Dunlop Street, Chevy Chase, MD 20740
(301) 405-2347 (W) (301) 656-0497 (H))
jschmidt@umd.edu

PROFESSIONAL PREPARATION

Allegheny College	English and Psychology	B.A., 1975
The Ohio State University	College Student Personnel Work	M.A., 1977
University of Minnesota	Counseling and Educational Psychology	Ph.D., 1983

APPOINTMENTS AT THE UNIVERSITY OF MARYLAND

Director, Interdisciplinary Research *The College of Education*	2003-Present
Director, Engineering Student Research *J. Clark School of Engineering*	1998-2003
Director, *Office of Institutional Research*	1995-1998
Assistant to the Vice President for Research	1983-1994
Adjunct Assistant Professor, Counseling and Personnel Services, *College of Education*	1985-Present
Licensed Psychologist in MD	1989-Present

PROFESSIONAL PUBLICATIONS AND PRESENTATIONS

Vogt, K., Smith, P., Schmidt, J., and L. Schmidt,(2004) "RISE: A Summer Experience for Incoming First Year STEM Students," presentation and paper, *Proceedings of the Women in Engineering Program Advocates Network (WEPAN) Annual Conference, Chicago, Ill*

Brown Leonard, J., Schmidt, J., Smith, P. E., and Schmidt, L., (2004) "A Pilot Investigation of Functional Roles on Engineering Student Teams", presentation and paper, *Proceedings of the 2004 American Society for Engineering Education Annual Conference and Exposition*, June 20-23, Salt Lake City, UT, *Nominated for Best Paper Award.*

Smith, P. E., Schmidt, J., Schmidt, L., and Vogt, K. E. "Research Internships in Science and Engineering (RISE): Summer Research Teams-Faculty and Students Benefiting from Role Model Hierarchies," presentation and paper, *Proceedings of the 2004 American Society for Engineering Education Annual Conference and Exposition*, June 20-23, Salt Lake City, UT, (2004).

Lent, R., Schmidt, J., Schmidt, L, Gloster, C., and Mourings, S, (2004) "Relation of Collective Efficacy Beliefs to Group Cohesion and Performance in Student Project Teams," presentation and paper, *Proceedings of the 2004 American Society for Engineering Education Annual Conference and Exposition*, June 20-23, Salt Lake City, UT

Schmidt, J., Schmidt, L., Smith, P. and K. Vogt, (2003)"Innovative Educational Opportunities for Women in STEM: Research Internships in Science and Engineering (RISE)", *Proceedings 33rd ASEE/IEEE Frontiers in Education Conference,* November 5-8, Boulder, Co.

Schmidt, J., Fines, J., Pertmer, G. (2003) The Teaching Fellows Program: Undergraduate Partners in Teaching. Paper presented at the ASEE conference, Nashville, TN.

Schmidt, J., Schmidt, L., Bigio, D., Smith, P. Engineering Students and Training in Teamwork: How Effective? Paper presented at the ASEE conference Nashville, TN.
Lent, R., Schmidt, J., Schmidt, L., Gloster, C. & Wilkins, (2003)G. Predicting the Academic Engagement of Women and Minority Students at Historically Black Institutions. Paper presented at the ASEE conference, Nashville, TN.

Anaine, S., Pertmer, G., and Schmidt, J. (2002) Keeping the Flame Alive: What Happens AFTER the ABET visit. Paper presented at the meeting of the American Society for Engineering Education (ASEE), Montréal, CA.

Smith, P., Schmidt, J., and Schmidt, L. (2002) Research Internships in Science and Engineering. Proceedings of the Women in Engineering Program Advocates Network Conference, Puerto Rico.

Dilli, Z., Goldsman, N., Schmidt, J., Harper, L., Marcus, S. (2002) A New Pedagogy in Electrical and Computer Engineering: An Experiential and Conceptual Approach. Paper presented at the ASEE/IEEE Frontiers in Education Conference, Boston, MA.

Lent, R.W. Schmidt, L., Pertmer, G., & Schmidt, J. (2002). Exploration of collective efficacy beliefs in student project teams: Implications for student and team outcomes. Paper presented at the meeting of the American Society for Engineering Education (ASEE), Montreal, Quebec.

SYNERGISTIC ACTIVITIES

(1) BESTEAMS: co-PI in BESTEAMS research project funded by Course, Curriculum and Laboratory Program of NSF (2001); organized three national conference presentations, multiple day-long dissemination and faculty training workshops, developed curriculum modules and initiated new research thrust on SCCT within the auspices of the College of Engineering.

(2) RISE: co-PI on project entitled: Research Internships in Science and Engineering (RISE) by the Program of Gender Equity of NSF (2001). Responsible for training predominately female faculty and student research teams in teamwork and mentoring skills, program assessment, and dissemination activities. Co-PI on REU site supplement to extend RISE.

(3) ASA: co-PI on NSF Assessing Student Achievement (ASA) program project entitled: "A Pilot Investigation of Functional Roles on Engineering Student Teams" (2003). Project involves conducting student focus groups to better understand how students effectively learn in teams.

COLLABORATORS & OTHER AFFILIATIONS

Collaborators: Robert Lent (UM), Patricia Mead (Norfolk State), Linda Schmidt (UM), Paige Smith (UM), Jeff Froyd (Texas A & M), Thomas Regan (UM), William Destler (UM), Paige Smith (UM), Jack McGourty (Columbia), Roger Azevedo (UM), Dennis Kivlighan, (UM), Chan Dayton (UM), Jean Dreher (UM)

Graduate and Postdoctoral Affiliations: Dr. Clyde A. Parker (University of Minnesota, retired); Ph.D. advisor; Dr. Robert Silverman (Ohio State University); M.A. advisor

Thesis Advisor and Postgraduate-Scholar Sponsor: John Fox (M.A. 1995, University of Arizona); John Zacker (Ph.D., UM, 1997); Terry Zacker (Ph.D., UM, 2001)

BUDGET

I.	Personnel A. Principal Investigator: Rachelle Brooks, Associate Director, Center for Assessment of Higher Education (1 month)	$ 0
	B. Co-PI: Dennis Kivlighan, Chair, Dept. of Counseling and Personnel Services (faculty member at institution of higher education; research conducted within term of appointment)	$ 0
	C. Co-PI: Janet Schmidt, Director of Interdisciplinary Research (1 month)	$ 0
	D. Graduate Research Assistant (part-time for a full academic year)	$ 6,022
	Total Personnel	**$ 6,022**
II.	Fringe A. Graduate RA + Tuition Remission	$ 5,800
	Total Fringe	**$ 5,800**
III.	Travel A. Travel to AIR Forum 2006 for PI	$ 1,500
	B. Travel to ASHE Conference 2006 for PI	$ 1,500
	Total Travel	**$ 3,000**
IV.	Participant Support A. Incentives	$ 8,000
	Total Participant Support	**$ 8,000**
V.	Other Direct Costs A. Computing Services	$ 7,000
	B. Two graders for essays	$ 4,800
	Total Other Direct Costs	**$11,800**
	Project Total	**$29,822**

BUDGET JUSTIFICATION

I. Personnel:

A. No funding is being requested for *Dr. Brooks'* salary for the one calendar month.

B. *Dr. Kivlighan* is a faculty member at the University of Maryland. This research is being conducted within his term of appointment.

C. No funding is being requested for *Dr. Schmidt's* salary for the one calendar month.

D. Funding for a part-time *graduate research assistant* is requested for a full academic year.

II. Fringe:

The proposing institution's usual accounting practices provide that its contributions to employee benefits be treated as direct costs. Fringe benefits provided for the RA for working part-time of a full academic year equal $2,000 and graduate tuition remission equals $3,800.

III. Travel:

A. *Travel funds* are requested to offset the cost of attending *the 2006 Association for Institutional Research annual meeting (AIR Forum)* by the Principal Investigator. Attendance at this meeting will allow the PI to disseminate results from the proposed research through formal presentation of results as well as informal exchanges during the meeting. Travel costs include round-trip airfare, transportation to and from the airport, registration, hotel accommodations and meals in accordance with the University of Maryland procedures.

B. *Travel funds* are requested to offset the cost of attending *the 2006 Association for the Study of Higher Education (ASHE) annual meeting* by the Principal Investigator. Attendance at this meeting will allow the PI to disseminate results through formal presentation of results, to investigate further avenues of disseminations (i.e. publications) and to develop future research based on initial findings. Travel costs include round-trip airfare, transportation to and from the airport, registration, hotel accommodations and meals in accordance with University of Maryland procedures.

IV. Participant Support:

A. *Incentives* will be paid to study participants: (100 participants @ $75) + (1 domestic airline ticket @ $500).

IV. Other Direct Costs:

A. The *computing services* of MIT Web Survey Service will be used for the survey's web design, programming and hosting.

CURRENT AND PENDING SUPPORT

Rachelle L. Brooks

- PI for AERA grant; 1 month/year (current).
- PI for NSF/ASA grant; 3 months – year 1, 1 month – year 2 (pending).

Dennis M. Kivlighan

- PI for Rehabilitation Training Grant from the US Department of Education; 1 month/ year (current).
- Co-PI for NSF/ASA grant; 0 months/year (pending).

Janet A. Schmidt

- PI for NSF REU program; 1 month/year (current).
- PI on NSF ITEST grant; 1 month/year – 3 years (pending).
- Co-PI for NSF/ASA grant; 1month/year (pending).

FACILITIES, EQUIPMENT AND OTHER RESOURCES

The proposed research will be conducted using the resources currently available to the Center for Assessment of Higher Education at the University of Maryland, College Park. These resources include office space, computers, telephones and other standard office equipment needed to carry out and disseminate the proposed research.