

WRITING A GRANT PROPOSAL

Malcolm Gordon, Ph.D.

National Institute of Mental Health

July 1999

INTRODUCTION

This technical assistance paper describes the process of applying for a research grant from the National Institute of Mental Health, how such applications are processed and evaluated, and the criteria used to evaluate grant applications. The author is responsible for programs that support mental health research in the areas of child and adult psychopathology. Most of the examples are drawn from these areas and the methodological and conceptual advice offered may not be applicable to research in other areas of research, such as biological research, basic behavioral research, and services research.

Grant application procedures. The National Institute of Mental Health is one of the institutes in the National Institutes of Health. NIMH funds grants for both individual research projects and research training/research career development. Grants for research projects fund one or more specific studies that aim to address an issue of relevance to understanding mental health and mental disorders. Research training/career development grants provides salary, stipend, and other support to train a particular researcher in research methodology in a specific mental health-relevant research area or to support an existing program at a particular university or research institution that trains researchers in research skills relevant to mental health research. A variety of both research awards and research training/career development grants are available. Types of research grants include regular research grants (referred to as "R01s" in the identifying codes assigned to submitted applications); small research grants (referred to as "R03s"); psychosocial treatment development grants (designated as "R21s"); and Multi-Institutional Collaborative Research (designated as "R10s"). Research training/career development grants include pre- and post-doctoral research training fellowships ("F31s" and "F32s"); Mentored Research Scientist Development Awards ("K01s"; "K08s" and "K23s") and Independent Scientist Awards ("K02s" and "K24s"). The different types of awards have various eligibility and budget restrictions. Brief descriptions of these awards are available on the Internet at www.nimh.nih.gov/grants/index.htm under "General Information About NIMH Grants, Programs and Program Staff Contacts."

Applying for a research grant is an involved and demanding process. All research project, career development, and institutional research training grant applications submitted to Public Health Service agencies, including the National Institutes of Health components, are submitted using the **Public Health Service 398** grant application form (State and local government units may use another form, but commonly also use the 398 form). Individual pre- and post-doctoral research fellowships are submitted using the **Public Health Service 416** application form. These application forms may be obtained from your college, university or medical school office of sponsored research, from the Office of Grants Information, Center for Scientific Review, National Institutes of Health, Suite 1040, 6701 Rockledge Drive MSC 7710, Bethesda, Maryland 20892-7710 (phone: (301) 435-0714) or directly from the Internet at www.nih.gov/grants/funding/funding.htm under "NIH Application Instructions and Forms." To download PHS 398 instructions and application form sheets you will need to use Adobe Reader software, but both the NIH PHS 398 and the PHS 416 webpages provides a link to download this software if you don't already have it.

NIH does not award grants to individuals, rather all grants are awarded to institutions that are responsible for the administration, especially the financial administration, of the grant award. Grant applications must be approved by and signed by a responsible institutional official. A research application is usually prepared by and submitted by a researcher who is designated as the "Principal Investigator" for the project and is responsible for the scientific conduct of the project. Many projects have more than one scientist with more or less equal responsibility for the project who may be designated as "Co-PIs" on the research team, but NIH does not formally recognize "Co-Principal Investigators." A single individual (the PI) is regarded as solely responsible for the scientific conduct of the

study. And although a research grant is awarded to the institution rather than the PI, the PI is regarded by NIH as key to the conduct of the research project. Thus, if the Principal Investigator moves to another institution, the research grant will usually be transferred to the investigator's new institution, provided it is feasible to conduct the research at the new site. Replacing the designated Principal Investigator (or other key personnel) on a research grant requires approval from NIMH.

NIH identifies areas of research in which research grant support is available in two ways: with standing **Program Announcements (PAs)** and specific **Requests for Applications (RFAs)**. Program announcements indicate the areas of research in which grant awarding programs at NIH typically award research grants. Most PAs indicate a range of relatively non-specific research topics in which researchers are free to develop research proposals (e.g., "Research on Co-Morbidity of Mental Disorders"). Requests for Applications typically identify specific research topics that programs are interested in supporting (e.g., "Research on Child Neglect"). RFAs usually target important, but under-researched areas, are often one-time or limited-in-time submissions, and often have a specific amount of funds set-aside to support the best applications on the research topic. Announcement of specific grant mechanisms and instructions for completing the grant application forms can be obtained from a number of sources. The NIH Office of Grants Information (phone: (301) 435-0714) can supply you with instructions for different types of award applications and announcements for specific requests for applications (RFAs). NIMH announcements and RFAs are available on the National Institute of Mental Health Internet home page at www.nimh.nih.gov/grants/index.htm. NIMH has also implemented an automated fax retrieval system to allow you to immediately obtain NIMH grant and program announcements by fax (Mental Health FAX4U). To access this system call (301) 443-5158 from a touch-tone fax machine and follow the voice instruction. A list of documents available on the Mental Health FAX4U system can also be accessed on the Internet at www.nimh.nih.gov/research/faxlist.pdf (Adobe Reader pdf format).

It is important - and it is the applicant's responsibility - to follow the instructions in the PHS 398 instructions and the specific program announcement for completing an application for a specific type of award. The instructions provide directions on what to include in the various sections of the application form, page limitations, deadline dates for submission of types of applications, eligibility requirements, budget and time limitations, and restrictions on materials that can be submitted (e.g., in appendices). In addition, changes in grant announcements, instructions, and application procedures are published in the NIH Guide for Grants and Contracts and, when they are published, they then become official NIH policy superceding the current published announcements. Therefore, you need to keep informed of such announced changes relevant to grant mechanisms that you might be applying for. Your institution's office of sponsored research should monitor such announcements published in the NIH Guide. You should

check with this office regarding possible changes in instructions or program announcements that supercede the announcements you might have. Alternatively, you can search the NIH Guide through the Internet for relevant research grant policy changes at <http://www.nih.gov/grants/guide>. Search the Guide using the title of the relevant grant announcement or related terms.

Completed research grant applications are initially submitted to the Center for Scientific Review (CSR, formally known as the Division of Research Grants (DRG)) of the National Institutes of Health. Currently, there are three deadlines a year for submission of regular research grant applications (and also Small Grants, and K Awards, see below). These dates are February 1, June 1, and October 1. Some other types of awards have different submission dates and you should consult the specific announcement or instructions for other awards.

Electronic Resources

It is becoming increasingly necessary for researchers to be facile at accessing electronic information, especially using the World Wide Web, in order to develop an NIH application and to conduct research. The following is a list of important websites that you should be familiar with in order to develop an NIH research application. Most of these websites will be referred to in the text:

www.nih.gov the NIH homepage

www.nimh.nih.gov the NIMH homepage

www.nih.csr.gov the NIH Center for Scientific Review homepage

www.nih.gov/grants/oprr/oprr.htm Office for Protection from Research Risks homepage

www.nih.gov/grants/forms.htm site to obtain grant application forms and instructions electronically

www.nimh.nih.gov/grants/granpol.htm site with announcements of latest policy changes affecting research grants

www.nimh.nih.gov/grants/pamenu.htm NIMH research program announcements

www.nih.gov/training/extramural.htm research training and career development announcements

www.csr.nih.gov/review/irgdesc.htm descriptions of NIH review committees

www.csr.nih.gov/committees/rosterindex.asp rosters of NIH review committees

www-commons.cit.nih.gov/crisp searchable data base of research projects funded by NIH and other Public Health agencies

www.nih.gov/grants/guide/index.html searchable electronic form of the NIH Guide for Grants and Contracts

Grant review. The NIH grant review and award process has two characteristics that distinguish it from some other Federal research agencies: Most NIH research grants are *investigator initiated* and NIH utilizes an independent *peer review* process to review grant application submissions. Because most research grant applications are submitted under very broad program announcements, a researcher is able to develop his/her own idea for an important research topic into a research proposal which can be submitted to NIH and evaluated for funding, provided it falls within the general research mission of an NIH institute (e.g., for NIMH, research on mental health and mental disorders or basic biological and behavioral research relevant to understanding mental health and mental disorders). Thus, NIH allows researchers in the field to determine research directions rather than having research directions determined solely by NIH staff. Almost all research and training/career development grants that are submitted to NIH are evaluated by standing review committees composed of reputable (in general, non-Federal) research scientists who evaluate the importance and scientific sophistication of proposed research projects. These standing committees are independent of the program components in the NIH institutes that actually award research grants. NIH staff from the grant awarding components are free to attend sessions of the review committees, but do not participate in or influence the review. The evaluation of the independent review committees as to the scientific merit of a proposed research project is, in general, the most important factor in whether a grant application is awarded.

All grant application to NIH are submitted to and initially processed by the NIH Center for Scientific Review (CSR, formally known as the Division of Research Grants (DRG)). Within CSR a submitted grant application is initially screened for referral to an appropriate Initial Review Group (IRG) for peer review and to an NIH institute or center for funding consideration.

Within six to eight weeks after submission of a proposal, the applicant will receive a mailed notification of the Initial Review Group to which the application has been assigned for review and the program within an NIH institute to which the application has been assigned for funding consideration. For further information about the Center for Scientific Review see their homepage at www.csr.nih.gov and for a description of application receipt, referral, and review see their webpage www.csr.nih.gov/review/peerrev.htm.

Most applications are reviewed by NIH review committees, although some applications are reviewed by special committees within the institutes (e.g., most RFA applications).

An IRG, also referred to as a "Study Section," is composed of reputable scientists working in various health and mental health-relevant research areas. They review the scientific merits of proposals. Proposals may be referred to a particular IRG either on the basis of the subject matter or of the methodology of the proposed research. A description of the areas of science reviewed by NIH IRGs is available at www.csr.nih.gov/review/irgdesc.htm. Members of NIH review committees are nominated and the review sessions are run by a Scientific Review Administrator from the Center for Scientific Review. Staff from the funding programs cannot participate actively in the deliberations of review committees, but program staff do attend, as observers, reviews of grant applications assigned to their programs and can often provide additional feedback on the discussion of a grant application by a review committee.

It is often helpful to access the membership list of a review committee that is likely to review your proposal for a number of reasons: (1) you should run a computerized literature search and review publications of members on the committee who conduct research in your area, as often reviewers use their own research experience or theoretical biases as a basis for critiquing proposals (if you do cite research results of committee members make sure you do not misinterpret or misconstrue their research results!); (2) you can be alerted to the level of expertise of the committee in the substantive or methodological issues important to your research and, if such expertise is lacking, raise the issue of outside reviewers with the Scientific Review Administrator; and (3) you can be alerted to possible personal biases or conflicts with committee members which might warrant recusing a particular reviewer or review by another committee. Rosters for NIH review committees are accessible via the Internet at www.csr.nih.gov/committees/rosterindex.asp. If you feel that the review committee assigned to review your application is inappropriate, contains members scientifically biased against your research approach, or lacks sufficient expertise to adequately review the science of your project, you should raise this with the Scientific Review Administrator of the review committee to which your application is assigned. You can request assignment of the application to another review committee or recruitment of outside reviewers to enhance the scientific expertise of the review committee in areas important to your proposed research. Such changes in review committee assignment or recruitment of additional reviewers is solely at the discretion of the Scientific Review Administrator. If through accessing the description of review committees from the CSR website and/or the rosters of review committees, you think that a particular review committee is most appropriate to review your application, you can request assignment of your application to a particular review committee in a cover letter that accompanies your application.

Currently, an application can either be approved and assigned a numerical Priority Score by a review committee, based on a consensus rating of the scientific merit of the proposal or not assigned a Priority Score, if the consensus of the committee is that the particular proposal does not fall within the top half of proposals in terms of scientific merit. A

mailer indicating the Priority Score received in the review is sent to the applicant usually within a week or two following the review meeting. The Scientific Review Administrator will also convey to you the results of the review group's evaluation of your proposal in a written "Summary Statement," that you should receive within 4 to 6 weeks after the review meeting. The Summary Statement consists of a synopsis of the committee's overall evaluation of the proposal and the original written critiques of the primary and secondary reviewers.

Priority Scores range from 100 (the best score) to 500 (the worst score). Applications are assigned a percentile score based on the distribution of Priority Scores for the committee over recent review sessions, such that the lower the percentile score the better the ranking of an application in the distribution of Priority Scores (i.e., an application with a 5 percentile is in the top (best) 5 per cent of distribution of Priority Scores for an IRG and a 95 percentile is in the bottom 5 per cent (worst) of the distribution). Applications do not receive a Priority Score if all the reviewers of the proposal agree that a proposal would probably achieve a Priority Score above the 50th percentile of the distribution of Priority Scores. This is indicated by a "***" next to "IRG Action" on the Summary Statement of the review of the application. Such applications are not reviewed by the whole committee. This is due to the large number of applications submitted and, since an application that receives a percentile score above 50% has almost no chance of being funded, reviewing only the best half of the applications allows the committee to focus its attention at the review meeting on applications which have some chance of being funded.

If your application does not receive a Priority Score, you should not necessarily assume that your research idea is not worth pursuing. You will still receive feedback in the form of written critiques from the assigned reviewers and can respond to the criticisms in a resubmission. Most research proposals must be revised one or more times before they score well enough to be eligible for funding. You should carefully read the reviewers' critiques of your proposal and consult with program staff about how "fatal" versus "addressable" the major criticisms of your study are. If you are able to revise and adequately address the major criticisms that resulted in the application not being scored, your revised proposal is likely to be competitive.

Reviewers are expected to use the following criteria in evaluating a research grant application: (1) significance of the proposed research; (2) the scientific sophistication and adequacy of the conceptual and methodological approach used in the research; (3) the innovativeness of the proposed research; (4) the qualifications, experience, and expertise of the researcher or research team to conduct the proposed research; and (5) the quality and adequacy of the scientific research environment to conduct the proposed research. Reviewers will still assign a single, global score to an application that is supposed to reflect the reviewers' overall evaluation of the "overall impact that the project could have on the <scientific research> field." Reviewers are expected to comment on each of the above five criteria in their review and to consider them in assigning an overall score, but their weighting of each criteria can vary from application to application. So, for example, a research project may not be rated very high on innovativeness, but still receive a good evaluation based on its significance for the field and the sophistication of the scientific

approach. It is hoped that this will result in funding of more research which is likely to have significant impact rather than research that is just technically well-developed.

Further information on peer review policy and criteria see www.csr.nih.gov/review/policy.htm.

An application's percentile score is one of the most important, if not the most important, criteria in funding decisions. Therefore, the impression that a proposal makes on the research experts composing the review committee is often the most important factor in a grant proposal being funded. Most proposals are not funded on initial application and must be revised one or more times before receiving a fundable priority score, if funded at all.

Before an application is eligible for funding it must pass through a second level of review by the National Mental Health Advisory Council, which meets approximately 3 months after the IRG meeting. In almost all cases the National Advisory Mental Health Council concurs with the IRG recommendation regarding review of a grant.

Research programs and program staff responsibility. Grant applications having direct mental health relevance or proposing to study basic behavioral, psychological, social, or biological processes of relevance to mental health are assigned for funding consideration by the Center for Scientific Review to the National Institute of Mental Health. Within NIMH research application will be assigned to specific programs, such as the Developmental Psychopathology and Prevention Research Branch (child psychopathology and developmental risk) or the Adult Psychopathology and Prevention Research Branch (adult psychopathology and risk), for potential funding and monitoring after funding.

Currently, research programs at NIMH are organized into three divisions. The Division of Basic and Clinical Neuroscience Research supports research that primarily focuses on basic and applied neurosciences. The Division of Services and Intervention Research supports research on mental health services and on interventions for child and adult mental health problems and disorders. The Division of Mental Disorders, Behavioral Research and AIDS supports research on basic and applied cognitive, social and behavioral processes, on adult and child psychopathology, and on mental health aspects of Acquired Immune Deficiency Disorder. Within each division are numerous branches and programs that focus on specific research areas, for example, the Developmental Psychopathology and Prevention Research Branch in the Division of Mental Disorders, Behavioral Research and AIDS is responsible for research on child psychopathology and within the Developmental Psychopathology and Prevention Research Branch is a program dedicated to research on child and adolescent trauma and victimization. An organizational chart for NIMH is available at www.nimh.nih.gov/about/compon.htm and a listing of program branch contacts is available at www.nimh.nih.gov/grants/pcc.htm.

The role of the program staff in the application process is to provide technical assistance to researchers seeking grant funding, to establish funding priorities in their substantive research areas, to request funding of approved projects, and to monitor grants funded in

this area. Usually questions about eligibility requirements for particular types of awards, questions about particular instructions for applications, and about materials that may be submitted with an application are best addressed to a Scientific Review Administrator in the Division of Extramural Activities. Substantive questions about the research itself and funding priorities should be addressed to program staff.

Grant application review criteria. In general a review committee will be concerned with four aspects of the proposal: its **importance** (e.g., Is it worthwhile studying what the proposal plans to study?, Does the research have implications for understanding, preventing, or treating important mental health problems? Will the field learn anything of value from the study? Will this study advance methodology or conceptualization in the field?); its **feasibility** (e.g., Is it likely that the researcher can accomplish the proposed research?, Does the researcher have the necessary expertise and prior experience to accomplish the work?, Will the researcher be able to recruit enough subjects?); the **technical merits** of the research methodology of the study (e.g., Is the conceptual basis of the research well-developed, empirically sound and likely to enhance understanding?, Is the researcher using the best measures available?, Is the proposed data analysis technically sound and does the researcher have the technical expertise to perform the analysis?, Has the researcher incorporated the important prior research in the area into the design, conceptualization, and analysis of the research?); and its **innovativeness** (e.g., Does the research break new conceptual or methodological ground in its scientific field? Will the knowledge gained be very different from existing knowledge in the field? Has similar research already been conducted in the field?).

Some more specific suggestions might be helpful:

Proposals that are Not Scored or that do not receive a good priority score usually fall into four categories:

- o Those that have a "fatal flaw" (i.e., an error in design or feasibility such that it is unlikely that the researcher will be able to draw valid conclusions from the data). "Fatal flaws" can take many different forms and can occur in (a) the sampling plan of the study, such as inappropriate or confounded study and control groups (e.g., comparing an identified clinical group with a community group without screening the community group for the clinical condition), or the infeasibility of obtaining the number of subjects proposed in the study; (b) the measurement plan of the study (e.g., an invalid measurement procedure for a key independent or dependent variable); or (c) the analysis or interpretation of the results (e.g., cell sizes too small to have adequate statistical power to detect group differences).

- o Those that are "incomplete" (i.e., not enough detail is provided on the measurement of variables, the way data will be analyzed, or the

conceptual framework or rationale guiding the research for reviewers to gauge the potential quality of the results). One common instance of this is a "trust me" approach to important aspects of the research (i.e., the claim that the proposer will be able to develop an adequate instrument or to get the sample proposed in the research without providing any documentation that this is likely to happen).

- o Those that have a serious human subjects concern (i.e., fail to describe procedures to adequately safeguard subjects from potential risks arising from participation in the research).

- O Those that are likely not to provide any new or important substantive knowledge or that test trivial, uninteresting, or overly general hypotheses

Proposals that are approved with a good priority score but which are not fundable usually have an adequate methodology, but have other problems:

- o Problems with feasibility (i.e., the scope of the research, the timeline for data collection and/or analysis, the lack of background of the researchers in the area of proposed research, or unacknowledged problems in obtaining or keeping adequate number of research participants make reviewers skeptical that the research can be conducted as proposed).

- o Problems with significance (i.e., the research is methodologically adequate but the problem being investigated is not of great significance, or does not appear likely to significantly advance the particular field of research, or tests a large number of disparate hypotheses without any differentiation as to the importance of individual hypotheses).

- o Problems with the "logic" of the research (i.e., the development of an adequate rationale for the relationship between hypothetical constructs and between constructs and measured variables. These types of problems are the most difficult to define and anticipate and include such instances as an inadequate account of why dependent and independent variables are expected to be related or the failure to consider important alternative variables or moderating variables).

One of the most common failings in applications that are not scored or scored with an unfundable priority score is a lack of clarity about important aspects of the research (e.g., a vague or overly-general conceptual scheme, no explanation of why certain variables are chosen to be measured, a sketchy indication of how study subjects will be chosen or

where they will come from, and a vague and quite nonspecific account of how data will be analyzed (e.g., a statement in the proposal's data analysis section that a large number of variables will be reduced or consolidated using "multivariate techniques"). Often the review committee will conclude that such instances of lack of clarity result from a failure of the applicant to carefully think through those aspects of the research.

Lack of clarity in proposals often stems from inclusion of too many variables in the conceptualization and design of the research. Problems in conceptualization, design, statistical analysis, statistical power, and interpretation increase geometrically with the number of variables in research studies. Variables to be included in the research should, probably, be divided into two groups, a relatively small number (say on the order of 5 to 10) of important key constructs and a remaining group of secondary or subsidiary variables. In the proposal you should be as clear as possible about how you hypothesize the key constructs are related, how you will define and measure them, and how you will statistically analyze them. Discussion of the secondary or subsidiary variables can be much less detailed. Applications that propose to collect data measuring a large number of variables, for example in a comparative study of rape victims and controls, to "see how the groups differ" often do poorly in peer review, especially, if there are no specific hypotheses about expected group differences or a discussion of how to reduce the number of variables to a manageable number for statistical analysis. Such a proliferation of variables often indicates a lack of thoughtful selectivity about what is really important to investigate in an area. One circumstance in which a larger number of variables can be incorporated into the research design is when you have a clear plan on how to reduce the number of variables by combining variables into more reliable or valid composite indicators of a smaller number of ("latent") constructs or by using your data to select a set of "best" indicator variables for these constructs. Unfortunately, many times applicants have only a vague or sketchy notion about how they will accomplish this (e.g., "the data will be 'factor analyzed'" or "LISREL modeling will be used.") It is often helpful to obtain expert statistical consultation on various data reduction techniques that may be suitable for your data. Be advised, though, that particular statistical manipulations of data may not make any conceptual sense for the kind of data you are collecting and you should understand enough about the particular techniques to judge the conceptual appropriateness of particular statistical techniques for your data.

Grant funding. Grant applications that are scored by an IRG and approved by the Advisory Council are eligible for funding consideration by the NIH institute to which the application has been assigned for funding. A research grant is an award from a Federal agency to an organization or institution to conduct research in the public interest; hence, funding a research grant application is solely at the discretion of the Federal agency.

Attaining a very high evaluation of scientific merit (Priority Score and percentile) from a review committee (IRG) does not guarantee funding of an application, although attaining such a high evaluation is one of the most important, if not the most important criteria, used in grant funding by NIH institutes. Advisory Councils and institute scientific and administrative staff can recommend applications be given both decreased and increased consideration for funding over and above IRG scoring. Decreased funding consideration can result from considerations of program portfolio

balance, budgetary impact, or significant scientific, feasibility, public policy, or ethical issues overlooked by the IRG or based on information unavailable to the IRG. Increased consideration can come from fit with program priorities or identification of a project as of high risk but with a very significant payoff in public health benefit, if successful.

Questions concerning application funding should be directed to the program staff person assigned as the Project Officer for your application. The Project Officer can recommend funding of eligible applications, but does not control this process, as funding recommendations are assembled and approved at the program, branch, division and institute levels. The funding process can be quite lengthy (a matter of months), particularly, if your application is borderline in terms of fundability. Your institution will be notified that your application has been funded through a Notice of Grant Award, which conveys the budget levels approved for funding of the application in the first and subsequent years, terms and conditions of the award, and where the funds have been deposited or can be accessed. It is the responsibility of your institution to notify you when they receive your Notice of Grant Award. You can consult with your Project Officer about your funding prospect, but program staff can make no commitment or guarantee of grant funding. The only way that NIH makes a commitment to fund a grant application is through the Notice of Grant Award.

THE PHS 398 GRANT APPLICATION FORM

Sections of the Application Form. The PHS 398 Grant Application Form has the following sections:

Face Page

Abstract Page

Budget Section

Biographical sketches of key personnel

Other Support section

Resources of the research site

Research Plan section

Assurances

Appendices

Follow the instructions that accompany the 398 form including limitations on type size, number of pages, and material that can be appended to the application. If you have questions about eligibility for different types of grants or how to complete specific items

or sections of the application, you should contact review staff in the NIMH Division of Extramural Activities or program staff in the NIMH's extramural programs, such as the Prevention, Early Intervention, and Epidemiology Research Branch.

It is particularly important that you consider the readability of an application in terms of its appearance (size and clarity of type), clarity of organization, and understandability. Reviewers, who must spend a great deal of time and effort reviewing applications before an IRG meeting, often comment negatively on applications that are hard to read because of illegible type or are confusing or hard to understand because of lack of clarity, contradiction between different parts of the application (e.g., use of different numbers for participants to be recruited in various parts of the application) or a confusing organization. Such negative reactions to the appearance and readability of an application may negatively impact on the evaluation of the application.

Some specific comments relevant to grant application sections follow:

Face Page. If you are applying for a specific type of grant, other than a regular research grant (an R01), such as an R03 (Small Grant), R21 (Developmental Psychosocial Treatment Grant), etc., you should indicate the title of the grant type and its announcement number on the second line of the face page to avoid a mis-specification of your application (i.e., having your application designated and reviewed as an R01 when you intended to apply for an R21). If you would like your application to be referred to a specific program at NIMH, you can also refer one of the program's Program Announcements on this line, if such an announcement exists (e.g., PA-94-094, Research on the Effectiveness of Children's Services for the Child and Adolescent Mental Health Services Program). Applications are, however, referred to review committees and programs based on the subject matter covered by respective review committees and programs and an application will not be referred to a branch that does not fund the proposed type of research simply because you request a particular program.

You must also enter information on both Human Subjects approval and/or Animal Welfare protection, if applicable. Human Subjects approval involves approval of the procedures for protection of human subjects in the research by a designated Institutional Review Board (IRB). This approval must be obtained prior to the review of the application and must be current within 12 months of funding of an application. If you have not obtained a final IRB human subjects approval prior to submission of the application or, if your institution does not have a Multiple Project Assurance agreement with the Federal Government, you should indicate "Pending" under the "IRB Approval Date." If your institution does have a Multiple Project Assurance, but your approval is not obtained until after you submit your application, you should contact the Scientific Review Administrator and inform him/her of the date that you obtained final IRB approval *before* the application is scheduled to be reviewed. If you are claiming an exemption from human subjects protection provisions, you should designate the exemption number in this section. Be advised, though, that NIMH interprets exemptions from human subjects protection very narrowly, and, if you intend to claim a human subjects exemption, you should consult with program staff concerning the acceptability

of such an exemption claim. For more information about human subjects protection, see the section on Human Subjects later in the discussion of the research plan.

Abstract Page. The abstract should contain a brief description of the background, goals, and methods of the research. Be aware that abstracts of funded applications are available to the public through NIH data systems that catalogue funded applications (e.g. CRISP) so it is best to write an abstract that would be understandable to an intelligent lay person and to avoid being overly "jargony" in the abstract. Moreover, funded applications are often monitored by various advocacy groups so it is prudent to include a sentence or two highlighting the potential mental health significance of the proposed research to avoid a perception that the research is not really relevant to "mental health" This is especially true of more basic research or research on mental health problems that might be viewed as not very significant (e.g., "interpersonal conflict," "marital adjustment" or "daily hassles"). Also avoid potentially inflammatory language, concepts, or claims in the abstract that might be much better explicated in the body of the text (e.g., "expressed emotion in schizophrenic families" or "sex offenders can be successfully treated in the community").

Budget. (see Modular grants below) NIH is attempting to streamline the grant application process and reduce the burden of preparing a grant application. This has resulted in some changes in the materials that must be submitted in an application prior to the review of the application, chiefly in the area of assurances and budget information. For some types of grant applications, detailed budget information need not be provided (see section on Modular Grants below) or supplied only prior to funding (see www.nih.gov/grants/guide/1996/96.05.17/notice-just-in-time-004.html for "Just-In-Time" procedures).

For those grant application that require detailed budget information, reviewers assess the project budget separately from the evaluation of the scientific merit of the proposed research. Nevertheless, questions about the budget can influence the evaluation of the overall merits of the proposal.

For most applied behavioral research, approximately 70%-80% of the budget will be personnel costs; therefore, estimating and justifying the number of staff positions and the amount of time for each staff position is the primary task in developing a budget. The budget will be examined by reviewers for its reasonableness in terms of the scope of work proposed; this is often performed on the basis of the reasonableness of cost per participant or cost per task.

The amount of the Principal Investigator's time budgeted to a project is usually in the range of 33% to 50% time for most of the field research projects we fund, but may range from 25% to 75% depending on the size of the staff and the complexity of the research. A typical time commitment seen in grant applications is 1/3 time during the academic year and 2 months full-time during the summer. If a Principal Investigator devotes much less than 1/3 time to a project, reviewers may think the Principal Investigator is indicating a lack of commitment to the research; if a Principal Investigator devotes much more than 50% time to a project, reviewers may think that the Principal Investigator is assuming

tasks in the research that are better left to (less expensive) support staff. Commitment of these lesser or greater percentages of the Investigator's time require justification. Some types of grants have requirements on the time commitment of Principal Investigators. There are also implicit norms in field research studies for some other staff positions. For example, statistical consultants are usually budgeted for a maximum of 10% to 20% time in the grant years in which most of the data analysis takes place and lesser amounts of time in the beginning years of the grant (unless the statistical analyst is actually a member of the research team performing other functions beside statistical consultation or analysis). Request for staff positions budgeted for time commitments much greater than normally seen require additional justification.

One can estimate the number of staff positions required to conduct the research by developing a flow chart for the various tasks in the research, including staff training and monitoring, the recruitment, assessment, and tracking of the research subjects, data coding and entry, statistical analysis, and writing up of results. When a project involves a large number of staff positions, it is especially important to justify each staff position by some indication of which activities each staff member will engage in and an indication of how much time such activities will require (e.g. "research assistant X will conduct an assessment with each of 100 subjects @ 4 hours per assessment and coding in addition he/she will ..."). Estimating number of staff positions and time per position will be more accurate if rates of subject recruitment, amount of time to completely assess all subjects in the study, time to locate previously recruited subjects in longitudinal research, and time to accomplish all proposed data analytic tasks are realistically estimated. Often a small pilot study in which subjects are recruited and assessed with the full research battery can both aid in the estimation of the time and resources needed to complete the research and to convince reviewers that the research can be accomplished in the time period and with the resources proposed.

There have been some changes in the interpretation of the allocation of administrative and clerical salaries to research grants (see, NIH Guide, Vol 23, Number 34, September 23, 1994). Formerly, it was quite common for research projects to include secretarial, administrative, and clerical support in the Personnel expenses of research grants. Current policy is that such secretarial, administrative, and clerical expenses are part of indirect cost reimbursement to the institution and may not be charged to the direct costs of the research except in special circumstances. Such circumstances might include situations involving an explicit agreement between your institution and DHHS that excludes secretarial help from indirect costs or documented, extensive research-specific tasks, such as, for example, the need to contact, recruit, or track a large number of research subjects for a study or to prepare extensive research materials used specifically in the research project (e.g., assessment forms or data entry forms). (Additional possible allowable clerical or administrative examples are cited in the above referenced NIH Guide announcement). If your research does involve extensive subject recruitment, subject tracking, data collection material preparation, data collection, or data preparation or entry that is specific to your research project, then personnel needed to perform such functions should, probably, be listed as "Research Assistants" rather than being specified as clerical

or secretarial positions. A careful specification of their research-specific functions should be given to avoid such personnel expenses being disallowed.

It is important that the budget timeline of the study be clear and reasonable in terms of subject processing and research tasks. For example, a review committee may wonder why a study requires 3 years to recruit and test 60 subjects or takes 2 years to analyze the data collected.

Often researchers who are submitting their first application skimp on their budget estimates on the (faulty) assumption that more economical the research budget the more likely the research will be funded. However, research with a \$70,000 budget is reviewed with the same criteria as research with a \$200,000 budget; namely, the scientific merit of the proposal. Although an exorbitant budget can create a negative impression on a reviewer, it is just as likely that an overly frugal budget can raise questions about the feasibility of the project, such as: Do the researchers really understand how much it is going to cost to get this kind of data? Moreover, if your project is funded, it will most likely be funded at the level you request or at a reduced level and it is extremely difficult to get this amount increased should your budget not be sufficient for the scope of work proposed. This is most likely to happen with complex projects and projects with hard to recruit participants (e.g. offenders, victims, and mental patients).

One common error in underbudgeting is failure to allocate staff time or resources to an important research task which is overlooked in developing a budget. For example, if the project involves longitudinal data collection and all staff time is computed on the basis of assessing subjects and no resources are allocated to tracking and contacting subjects, reviewers are likely to think that the project will not have sufficient resources to manage subject attrition - a major problem in longitudinal research - over the course of the project.

Other categories of expenses also need to be carefully estimated. Often the office of sponsored research at your institution or knowledgeable colleagues who have conducted similar studies can aid in estimating such expenses as secretarial and office expenses, rent, and subject recruitment expenses.

If you are proposing more than one study in the overall project, it is helpful to provide enough information so that the Committee can estimate the costs for each study within the overall project. This is because the Committee sometimes will approve funding for only one or some of the studies within the overall project and not others. When this occurs and if there is not enough information to estimate the budget costs of the reduced scope of the research, the Committee must guess how much of the budget to cut – possibly with insufficient funding for the approved studies. Also, if you already have other types of grant support or are applying for other support, you must clarify how your total time will be allocated among the various types of support that you receive (e.g., that your total support does not exceed 100% of your time). If you should receive additional support (e.g. from Federal programs, foundations, or private funding sources) prior to the actual award of funds or during the course of your project that would necessitate a

reallocation of your time to the project or that would change the source(s) of your salary support, you should consult with your government project officer with regard to implications for the project budget.

A few budget categories that are often overlooked are the following:

1. Participant fees - Such fees can often help in recruiting participants and in decreasing dropout rates, especially in longitudinal studies. With difficult-to-recruit participants the amount of the participant payment should be set high enough so that it is likely to serve these functions (i.e. as an incentive for participation and to reduce dropping out) rather than set at the lower level of an honorarium or to only cover expenses for participating in the research. The amount of participant payment should reflect the amount of time and effort required of participants. On the other hand, participant payment should not be so high as to be coercive for participation, for example, with low-income participants or children. Researchers who collect data on multiple occasions from the same participants often build in extra incentive payments for completion of all the data collection sessions to avoid incomplete data. However, it is unethical to withhold payment from participants for the data collection sessions they do attend.
2. Consulting fees - By allocating funds for subject matter and methodological experts this is often a relatively cost-efficient means for bolstering both the expertise of the project staff and the prospects that the review committee will believe the project is feasible. Complex projects may have an Advisory Board that meets at least annually to advise the project staff on design and conceptualization of the ongoing study.
3. Costs associated with participant tracking or record searches - For projects collecting longitudinal data or data from official records it will often be beneficial to either add additional staff solely for this purpose or to propose some mechanism (with associated expenses) to maintain contact with participants (e.g. periodic call backs) or to extract data from records.
4. Subcontracts - Subcontracts can be used to help overcome feasibility problems (e.g. a subcontract to a service delivery agency that might have access to subjects it would otherwise be difficult to recruit or to a survey research firm to collect survey data).

Modular grants

Beginning with the June 1, 1999 grant submission deadline research grant applications with (direct cost) budgets that do not exceed \$250,000 in any year of the proposed grant project period must be submitted as "modular grants." Research costs (direct costs) of the research must be proposed in increments of \$25,000 up to the \$250,000 maximum for modular grants. Modular grants incorporate a number of streamlining changes to the traditional grant application, especially with regard to budget documentation. For modular grants detailed budget for the first and succeeding years of the grant are not submitted nor are "Other Support" pages required on submission of the grant application (they may be requested later by

program staff if an application is to be funded). Specific salary information is not required but the number and percent of effort of key personnel is still required. It is expected that the modular grant levels will remain constant in each year of the grant. If this is not the case (for example, it is quite common for data collection years of a grant to be larger than a start-up or data analysis year of a grant) the increase in modular level in later years of a grant over the initial year of the grant must be justified. Changes association with grant submission under modular grants does not mean that you can not submit a grant application for a budget larger than \$250,000. Grants requiring a (direct costs) budget of more than \$250,000 in any year of the grant project must be submitted in the traditional form with detailed budgets. For more information on modular grant application with examples of modular grant application forms, see www.nih.gov/grants/funding/modular/modular.htm.

Budget limitations and reductions

If you have not had previous Federally funded research grants nor considerable experience in conducting large or expensive studies, it is usually wise to avoid requesting funding of a very large or expensive project for your first application. Research mechanisms with smaller budget limitations (such as B-Start, Small Grants, and Exploratory/Developmental Grants (R21) for Psychosocial Treatment Research) should be considered, if feasible in terms of the scope of necessary work and the budget limitations. Review committees are likely to hesitate to award a fundable priority score to an investigator without a track record in large-scale funded research.

Because of the large number of well-reviewed research projects eligible for funding by NIMH, the size of a project's budget has become a factor in funding decisions. Since funding projects with very large budgets will absorb a large percentage of a program's limited research budget and result in funding of a smaller number of additional grants, **NIMH now requires applicants with a project budget exceeding \$500,000 in direct costs in any year of the project to notify program staff of their intention to submit such a project and receive permission from NIMH to submit such a large grant application. Failure to so notify the relevant NIMH program prior to submission is likely to lead to the return of the application without being reviewed or considered for funding.** Although policy on large grants differs in the various NIMH programs, many programs are highly likely to impose significant budget reductions on such large grants, primarily, because large grants consume a large percentage of a program's budget and may preclude funding other good research projects equally deserving of support. Thus, if your research project has a very large budget, you should consult with program staff. I would recommend that you should carefully consider the possibility of breaking up the project into separately submitted smaller projects. Although this creates more work for the applicant, and is sometimes not feasible to do, failure to consider this possibility risks having your research budget significantly reduced, anyway, or not funded at all. This is especially true if your prospects of funding are somewhat borderline or if your project with a reduced budget lacks the scientific merit of the original proposed project.

NIH has also imposed some restrictions on year-to-year budget increases in funded grants and imposed other reductions in overall grant budgets. This has been necessary to continue to allow NIH to fund a reasonable percentage of meritorious applications. Currently, year-to-year increases in budget categories (i.e., Personnel, Equipment, etc.) are limited to an inflationary increase based on an estimate of the yearly increased cost of conducting biomedical research and has recently been in the range of 3 or 4%. Increases above this amount in each category have to be justified by an **increase in the scope of the work** (this is usually due to increases in, for example, number of subjects recruited or amount of time devoted to data analysis over the preceding budget period). Thus, if your institution is increasing salaries by 6% a year and you build this percentage increase into your Personnel budget levels from year to year, this percentage increase in budget levels (greater than the allowed 3% increase per year) will not be accepted by NIH. Either your institution will have to absorb the additional increase or you will have to economize in other ways. How the inflationary budget cap on category increases is actually applied is somewhat complicated and you should consult with program staff about this issue.

It has been usual for NIMH in the past few years to reduce budgets of new grants to ensure adequate funds to fund a reasonable number of new grants, so do not expect to have your research fully funded, but instead expect to have to negotiate new budget levels and possibly a reduction in the scope of the research to meet the new reduced budget levels

Biographical Sketch

Reviewers examine the listed experience and publications of the Principal Investigator(s) for evidence that he/she is productive and has the experience and expertise to conduct research in the proposed area. In this regard, research publications in the proposed area of research in major, peer-reviewed journals count much more heavily than do publications in minor journals, talks and presentations, and non-research papers. If you do not have such publication credits, it might improve your chances of approval to conduct at least some pilot research in the proposed area, which you can report in detail in the "Preliminary Studies" section, to demonstrate your research competence in this area.

Sometimes an applicant will submit an application to conduct research in a field related directly or indirectly to his/her own area of research expertise or using a type of methodology he/she hasn't used before (e.g., an applicant has conducted basic studies of the differences between abused and nonabused children and develops a proposal to develop and/or evaluate an intervention program based on findings from the previous basic research). The absence of direct research experience in the proposed type of research (in this case, absence of intervention research experience) will often be viewed as a weakness in the qualifications of the researcher to conduct the research. In such cases, I would strongly urge you to collaborate with an individual or team that does have considerable research experience in the type of research proposed.

Another issue that may arise with respect to the qualification of staff to conduct the research is the experience of the Principal Investigator in conducting research on the scale proposed in the application. For example, if an applicant's prior research has been cross-sectional studies with relatively moderate sized samples and the applicant is proposing a major longitudinal study with a large sample size, the review committee may question the ability of the applicant to deal with all the complexities of subject recruitment and retention and data collection in a large, longitudinal study, if the investigator has not had such experience. In such cases project staff or consultants with such experience may mitigate this concern to a certain extent, but consideration should be given to scaling back the size of the study.

Research Plan

The purpose of the Research Plan section of the application is to clearly present to review committee reviewers the details of why and how you propose to conduct a particular research study or related research studies. This includes a clear statement of the specific aims of the research, the rationale for why the study should be done, the conceptual framework that motivates the research questions and research design, and details of how the research data will be collected and analyzed. Reviewers are not interested in a general review of a research area or a general description of a program of research, but in the specific details of why and how a specific research project will be conducted. Many criticisms of reviewers concern issues of conceptualization, design, and methodology in which the applicant has failed to provide sufficient detail about aspects of the research (e.g., failure to clearly specify or define the key constructs or their relationships being investigated in the research, failure to make clear what measures are being used in the research or their relationship to constructs in the study's conceptual framework, or failure to clearly state the specific data analysis procedures that will be used to test research hypotheses).

It is difficult to provide all the details necessary to fully describe proposed research in the narrative portion of the application. Page limitations and restrictions on material in Appendices are enforced; attempts to circumvent the page and Appendices limitation (e.g. by printing the proposal in reduced type) may result in your application being returned to you without being reviewed. One good strategy to attempt to convey more information in less space is to pack summary information into charts or tables (e.g. a summary of studies of reliability and validity of the measures included in your research). Material can also be included in Appendices and referred to, provided it conforms to restrictions on the contents of Appendices. Questions about these restrictions should be addressed to the Scientific Review Administrator of the assigned Initial Review Group in the Center for Scientific Review.

The remaining sections of the paper discuss details of describing a proposed research project.

Specific Aims

The aims of the research are the specific research questions, hypotheses, or overall theory that the research is seeking to address or test. The aims of the research usually fall into the categories of (a) developing knowledge that will help to understand or explain a phenomenon (e.g., to assess the relative impact of hypothesized individual, familial, and social factors on outcomes of sexual abuse on children), (b) filling gaps in knowledge in the research field (e.g., although it has been shown that many delinquents desist in their serious delinquency, the factors that determine this desisting have not been identified), (c) to test an established or proposed model or theory of a phenomenon (e.g., a study to test learned helplessness theory in a sample of women who return to a battering relationship) or (d) to assess or test interventions or factors that are relevant to clinical intervention (e.g., a study to assess cognitions concerning use of physical force in sexual relationships among early adolescents as such cognitions may be relevant to early interventions to prevent perpetration of coercive sexual relationships). The aims of the research should be specific and clearly stated. Often the specific aims of the study are stated as hypotheses or specific questions to be addressed in the research. The number of specific aims should reflect the important hypotheses or questions driving the research rather than presented as a laundry list of research questions or hypotheses of varying degrees of importance. Material not directly relevant to the statement of the aims of the research should be placed in other sections of the applications, (e.g., background justification and references in the literature review section, data collection methodology in the Research Plan section).

The specific aims are extremely important in developing and organizing a research grant application, as the Background and Significant sections should lay the groundwork for the importance of the aims, prior research relevant to the specific aims, and the feasibility of accomplishing the aims and the research plan should be developed to actually be able to achieve the aims of research. Reviewers assess both the importance of the specific aims and the likelihood that the research plan will be able to achieve the specific aims (i.e., answer the research questions or adequately test research hypotheses or a theoretical framework).

Background and Significance

Review committees generally prefer research in areas that are clearly of great significance to the mental health research and service delivery fields and prefer proposals that provide a theoretical rationale for the research, rather than merely investigating relationships between measures in a research sample. This preference for theoretically motivated hypothesis testing or theory building research can be a double-edged sword, however, because if the committee feels that the theoretical rationale is not adequately developed or not cogently linked to the research questions or if they just don't like the proposed theoretical orientation, this can be detrimental in the evaluation of the proposal.

You should clearly state why it is important to your research field to do the research you propose and what is the mental health relevance of the research. NIMH funds research on mental health and mental disorders in adults, adolescents or children or basic and applied research that is relevant to understanding, preventing, or treating mental health problems.

Research studies must, centrally, include measurement of mental health, mental health problems, or mental disorders or of biological, psychological, behavioral or social factors or processes relevant to mental health. An adequate rationale for the mental health significance and for the particular conceptualization guiding the research should be presented in the review of previous research in the field. Sometimes a technically well-developed proposal receives only a fair priority score when the reviewers feel that the fundamental question underlying the study is not very important or interesting. It is very difficult to improve such an application enough to have a chance at being funded because reviewers are just not enthusiastic about the basic question underlying the research. Therefore, it is important to emphasize the importance of your research both in terms of its implications for research and theory construction in your area and for public mental health in general or for specific population subgroups. Your case for the importance of your proposed research should leave reviewers with the view that the results will make a difference to the research field and, particularly, have an impact on conceptualization and/or methodology in an important mental health area.

The committee usually assesses the adequacy of the literature review that serves as a justification for the proposed research, viz., are any important studies or areas not covered in the literature review? Hence, the research review should thoroughly cover the area of research and relate the proposed research questions to prior research in the area. Reviewers expect a thoughtful and critical evaluation of the extant research literature; that is, the literature should be more than a summary of what has been reported in other studies. In particular, an adequate foundation for your conceptualization and research hypotheses should be laid in the review of previous research. The literature review should provide a basis for the choice of concepts being investigated, the conceptual framework underlying the research, cite prior research with these concepts in the proposed study area and in related fields, and critique alternative conceptual frameworks. It need not be encyclopedic, but should weigh the evidence for and against the particular theoretical framework or the hypotheses you are proposing. In particular, research findings that may be interpreted to contradict your hypotheses or theoretical framework should be critically evaluated. Use of searchable electronic data bases of the research literature (e.g., Medline, Web of Science, Social Science Citation Index, or PsycINFO) is strongly recommended. These data bases should be searched for citations on previous research with the key constructs being investigated, the measures proposed for use in the research, and the methodology (e.g., statistical procedures) that are key to the research. A particularly valuable electronic resource for developing an application and accessing research trends and funding program priorities is to search the CRISP (Computer Retrieval of Information on Scientific Projects) data base at <https://www-commons.cit.nih.gov/crisp>. This searchable data base contains abstracts, investigator and funding information on biomedical research projects funded by NIH and other Federal agencies. Subject matter or other search terms can be entered to identify past and currently funded research projects. Searching the CRISP system is an excellent way to determine the current state of art of research in an area, funding priorities of Federal programs, and to identify researchers who are conducting sophisticated research in a particular subject matter area.

The literature review should be oriented toward the specific questions or hypothesis that your research is concerned with. One of the common weaknesses of literature reviews in proposals is to include a very general survey of a research area that provides no justification for the particular hypotheses or concepts that you put forward (e.g. a general discussion of the outcome literature on child abuse, but no focus on particular studies of increased aggressiveness - when that is your major hypothesized outcome variable). In particular this discussion **must** provide adequate justification for the two most important choices on conceptual issues made in the research: (a) **Why were the particular concepts being focused on in the research chosen, why are these concepts important (as opposed to other concepts)?** and (b) **What is the justification, in terms of previous research, pilot research, or a theoretical framework, for specific hypothesized relationships among constructs to be investigated in the research?**

Preliminary Studies

This can be one of the most important parts of the application, because it can convince reviewers that you are likely to achieve your research goals since you have previously completed relevant and technically competent work in the research area of the proposal. Prior research or pilot studies assessing the important constructs hypothesized to be related in the research, assessing the psychometric adequacy of new measures you intend to use in your research, or piloting interventions can be important in convincing the reviewers that the research is likely to achieve the specific aims driving the research. Enough detail of prior studies should be given to convince reviewers of your technical competence to conduct the research plan. **It is important to describe the results obtained from prior research that are relevant to the proposed research, including results that bear on the feasibility of the proposed research, on the adequacy of sample sizes you will have access to, and on the likelihood that the major hypotheses of the study will be supported.** As with the literature review, your discussion of your previous research results should focus on their relevance to how you conceptualized and planned the research proposed in the application, especially the specific hypotheses or conceptual framework guiding the research.

Research Design and Methods

In assessing the technical merits of a proposal a review committee considers 5 aspects of the proposed research: (1) the conceptual framework guiding the research, (2) the measurement of key variables, (3) the sampling plan of the research, (4) procedures for data collection, and (5) the data analysis plan for the research.

Conceptual Framework to Guide the Research

There are four types of proposals that are submitted: (a) exploratory studies, in which an investigator identifies a number of important variables and proposes to investigate empirically the relationship among these variables; (b) hypothesis testing studies, in

which the investigator, on the basis of the literature in the field or his own previous research, proposes to test specific hypotheses about relationships among variables in particular groups or samples; (c) model/theory building studies, which are similar to (b), but in which, in addition, the investigator proposes a theory or model that includes constructs and relationships among constructs and that accounts for important findings from previous research; the proposed research will further refine or elaborate on the proposed model/theory; and (d) model/theory confirmation or testing studies, in which the investigator has developed a relatively complex and conceptually elaborate theory (e.g. specifies relationships among constructs over time or between domains) and is testing the fit of the model to a data set or is comparing the fit of his model to an alternative model on a data set (e.g. with structural equation modeling).

In general the frequency of proposals received decreases down this ordering, but review committees evaluate the sophistication of proposals in this order (provided, of course, they think the model or theory is tenable and sufficiently developed and/or they like/agree with your theory). In particular reviewers often speak approvingly of a proposed research as having an adequate "conceptual framework" to guide the research. Most research proposals fall in the first category of "exploratory research"; such proposed research has a low probability of being funded. This does not mean that exploratory research is not fundable, for example, in areas in which there is a dearth of research or little or no theory development, but you have to write a very strong proposal in terms of state-of-the-art data collection and analysis to provide justification for doing solely exploratory research. Most exploratory research should be upgraded to at least hypothesis testing research by stating specific research hypotheses derived from previous research in the area or from the research literature in the area. In competing continuation proposals the review committee has a strong expectation that the research will clearly be at the hypothesis testing phase and preferably at the model building phase.

Thus, an important criterion review committees use in evaluating research proposals is the specification of an adequate "conceptual framework" to guide the research. **In fact the adequacy or sophistication of the conceptual framework is probably the one aspect of an application that distinguishes applications that receive priority scores that result in funding from very good applications that do not quite get funded.** This conceptual framework should be developed in the "Background and Significance" and/or "Preliminary Studies" sections of the application. Such a conceptual framework may consist of an established or postulated theory to explain the behavior or phenomenon being studied (e.g. the "differential bonding" theory of delinquency causation or Finkelhor and Brown's four factor traumagenic theory of the effects of child sexual abuse) or be a set of postulated relations between independent/antecedent (causal) variables and dependent/outcome variables based on your own or others' previous research and/or clinical experience. In general the priority score of your proposal will depend on how well you articulate and justify the conceptual framework underlying your proposed research and how well you use this framework to formulate specific hypotheses, to design the research study, to choose measures, and to specify a data analysis plan.

Development of a Conceptual Framework

Constructs

You must clearly specify the conceptual domains you intend to investigate (e.g. the relation between the domain of "family functioning" and the domain of "antisocial behavior in children").

The specific constructs within each domain (e.g. "parental acceptance/rejection" or "parental modelling of antisocial behavior" within the domain of "family functioning" and "firesetting behavior" or "physical assaults" in the domain of "antisocial behavior in childhood") should be clearly identified. You should pay particular attention to the conceptualization and operational definition of the most important independent and dependent variables in each of the conceptual domains. This is especially true with constructs that encompass individuals, behavior, or other phenomena with a variety of manifestations or that cover an extensive range on a continuum (e.g., "abused children," "aggressive behavior," or "self-esteem"). You should be aware of and cite the important literature in the field that has implications for the definition, differentiation, or subcategorization of broad general constructs (e.g., factor analytic studies of general scales measuring the global construct).

You should adequately identify which manifestations of constructs you are including in your research. Often when theory or empirical research becomes more sophisticated in a particular domain, specific constructs in the domain become more sharply defined or more differentiated and better defined operationally. Thus, your own conceptualization of a research domain may be viewed as more sophisticated if the constructs you define are better differentiated and/or better defined operationally than previous research in an area. For example, 5 to 10 years ago research that investigated differences between "abused" children and "nonabused" children, in which "abuse" was defined as "confirmed by a protective services agency," might be fundable. However, with growing sophistication in research and conceptualization, a review committee would now probably require a differentiation of the "abused" group into different types of abuse, (e.g., physical or sexual abuse) or a characterization of the severity of abuse, and more elaborate operationalization of "abuse" (e.g. independent confirmation by coding of investigative records or interview with a parent about the types of abusive incident(s)), and would also probably require more evidence to warrant accepting a child as "nonabused" than "absence of a protective service agency confirmation".

Often with constructs that represent categorical study populations (e.g., groups with DSM diagnoses) operationalizations of the categories defining the study populations may differ considerably in who is included in each category and their characteristics. You should clearly state the specific criteria that include and exclude subjects in the various study categories and how the inclusion and exclusion criteria are to be applied in selecting particular subjects. For example, "sexual abuse" can be variously defined, especially with

regard to type and severity of sexual acts involved (e.g., contact versus noncontact sexual acts, age difference of victim and perpetrator, amount of force or coercion accompanying the sexual acts) and can be assessed by various means or from various sources (e.g., various levels of confirmation from protective service agencies or courts, from self-reports or reports from others using various questions or interviews at varying levels of specificity). Complaints about fuzziness of constructs leading to confusion concerning who exactly is included in study populations and their characteristics is a recurrent criticism made by review committee of unfunded applications.

You should also consider the implications of your definitional criteria for categorical constructs on how these constructs can relate to other constructs in the conceptual framework of the study. Various conceptual and operational definitions of "sexual abuse" can have serious implications for relating this construct to other constructs in a conceptual framework. As an example: if the construct of "sex abuse" was operationally defined by means of recruitment of subjects referred to a treatment program for sexually abused children and adolescents, this may result in a very heterogeneous or a very restricted population with regard to abusive experiences depending on admission criteria and referral patterns (e.g., all cases may be referred or only the most symptomatic cases referred, males may not be referred). The resulting study population may have significant implication for the relationship between "sexual abuse" and other constructs (e.g., relationship with various symptoms, if cases are referred on the basis of seriousness of symptomatology, or the relationship between severity of abusive experience and other variables, if selection restricted the range of severity of abuse). Thus, you should be very clear about the operational definition of these types of important categorical constructs. In cases in which you don't control the identification of the cases (e.g., if you recruit "confirmed" cases of sexually abused children from protective service agencies) you should collect enough information on potential definitional characteristics of the cases (e.g., type of abusive acts, severity) to allow an examination of the study sample according to various definitional criteria).

A problem with increasing differentiation of a global construct is that it will multiply the number of groups or variables studied and, hence, the number of subjects required to attain adequate statistical power to compare groups. Thus, if 50 "abused" and 50 'nonabused' subjects are required to adequately test a hypothesis, differentiating the "abused" group into 3 subgroups triples the number of abused subjects required (who are also probably more difficult to recruit than controls) in order to adequately compare the abuse subgroups. Thus, your proposal requires a type of balancing act between differentiation of global constructs and number of subjects required to adequately analyze data from subgroups or additional measured variables.

Multivariate and Multicomponent Constructs

Many behavioral and psychic constructs have multiple components or multiple manifestations or are multidimensional. To give just a few examples:

"Depression" has emotional (e.g., anhedonia, sadness), cognitive (e.g., low sense of self-efficacy, hopelessness), and behavioral (e.g., low activity level, irritability) manifestations.

The diagnosis of Post Traumatic Stress Disorder includes manifestations of re-experiencing (e.g., flashbacks), hyperarousal (e.g., anxiety and fearfulness), withdrawal (e.g., avoidance of situations similar to those in which the trauma occurred), and hyperreactivity (e.g., increased autonomic reactivity).

Children may demonstrate different amounts and different forms of "aggression" in different settings (e.g., at home, in school, with peers).

"Self-esteem" may have different dimensions on which it is evaluated (e.g., social, physical, and academic competence).

It is not usually the case that the various manifestations, components, or dimensions of these complex constructs are perfectly correlated (at one extreme) or are independent (at the other extreme). The various components of different constructs can vary in the degree of relationship among the constituent components and, further, the pattern and degree of relationship among components can vary for different individuals or groups of individuals. For example, some studies have indicated that many individuals who are diagnosed with PTSD alternate in manifestation of intrusive and withdrawal symptoms; similarly some individuals may manifest much stronger cognitive than behavioral symptoms of depression. Many research studies find positive but rather modest correlations between the various components, dimensions, or manifestations of complex constructs and often find higher consistency among components or manifestations among individuals at the extremes of the distribution of the manifestation of the construct (i.e., more agreement between affective, cognitive and behavioral manifestations of depression among nondepressed or seriously depressed individuals).

How the constituent components, manifestations, or dimensions of complex constructs are handled has serious implications for all aspects of research conceptualization, design, and analysis, including the conceptual framework of the study (varying relationships between the components of the construct and other variables), the choice of measurement instruments (some instruments or procedures might measure some components of the construct better than others), the sampling plan of the study (subjects might be chosen based on expectations about the consistency or strength of relationship among the various components of the construct), the research procedures adopted for the study (data for contextually dependent behaviors might have to be collected in multiple settings or situations), data analysis of the study (individual measures of components or a composite measure of the construct may be the chosen for analyses), and the interpretation of results (the strength of relationships among variables may depend on which components of complex constructs are being related). Careful attention should be paid to the global and constituent components of complex multicomponent constructs in planning all phases of the research.

Relationships among Constructs.

A conceptual framework includes not only the identification and operationalization of constructs, but also the specification of the important interrelationships among the identified constructs. Behavioral research uses a heterogeneous collection of different types of relationships among constructs, including deterministically causal (i.e., one construct is necessary and sufficient for the occurrence of another construct), probabilistically causal (i.e., the occurrence of one construct increases or decreases the probability of another construct, such as "risk" or "protective" factors) or cooccurrence (i.e., occurrence of constructs is correlated with no specification of causal direction). You should specify the most important relationships among constructs either in terms of which constructs are known to be related from prior research, which relationships are most important in the hypotheses guiding your research, or are most important in terms of new or innovative ideas about how constructs are interrelated. Often conceptual frameworks only specify which constructs are related and which are not.

Research often extends conceptualization in a field by specifying more specifically the relationship among constructs (e.g., rather than global constructs being related, either subcategories of the constructs are related, such as only "physically abused" children rather than all "abused" children manifest externalizing behavior) or by adding an additional construct to a relationship net among constructs (e.g., identifying a cognitive mediator of particular symptoms in abused children). If you consider the importance of specifying relationships among constructs in a conceptual framework, you can see why it is necessary to limit, in the application, the number of constructs posited in the conceptual framework. The number of just **binary** relationships among constructs in a conceptual framework equals the number of constructs (n) times (n-1). Although it is not necessary to specify the relationship among all the different constructs in your conceptual framework, the more constructs you have the less likely you will be able to adequately specify and justify the various interrelationships and the more likely a reviewer is to object that you have not adequately described how the various constructs are interrelated.

Hypothesized relationships among constructs should be justified either from previous research, the deductive logic of the conceptual framework, or the literature review. In general a conceptual framework is considered as more "sophisticated" to the extent that it specifies "linkage" or "mediating" variables between "input" and "output" variables. For example a research proposal that specified and operationalized some cognitive or emotional processes that mediated the relationship between abuse characteristics in childhood and mental health effects on the child would be viewed as more "sophisticated" than a conceptual framework that just predicted empirical relations between abuse characteristics and mental health symptoms. The presentation of the conceptual framework (i.e. conceptual domains, specific constructs, and hypothesized relationships among constructs) in the application can take many forms, such as a verbal account, path diagrams, Venn diagrams, regression equations, or structural equation models.

You should consider possible problems with the proposed conceptualization (e.g. alternative conceptualizations of the specific constructs in a domain, alternative hypotheses about the relationships among the constructs, and important moderating or confounding variables you have omitted or overlooked). Attention should also be given to the temporal relationship among constructs. For example, if you posit that social support decreases the probability of (mediates) depressive symptoms in women sexually abused in childhood, it makes a difference whether you are positing that this relationship is temporally true of social support **at the time of the abuse** (in which case you need some adequate assessment of this construct at that time), **ongoing** since the abuse, or **at the time you are assessing depressive symptoms in adulthood**.

The length of time between occurrence or assessment of related constructs is also of importance. Temporally distant input and output variables are harder to justify as being linked than are contemporaneous or temporally contiguous variables. For example, attempting to relate childhood variables (e.g., abuse experiences) to adult outcomes may be criticized as too temporally distant to indicate a direct effect, especially if intervening variables are not specified and assessed. Although a model of (early) exposure being related to group differences in exposed and nonexposed groups is often accepted in epidemiological research, it is less readily accepted by social and behavioral researchers who tend to favor mechanistic models (i.e., exposure -> mediating variables -> outcome). Therefore, if you do propose a model relating temporally distant variables, you should include some discussion of the validity of inferring a causal relationship between the variables across the time period.

Hypotheses

The relationships among the specific constructs that you specify in your conceptual framework are the basis for the formulation of the research hypotheses guiding the research. For example, if your conceptual framework includes such global and specific constructs as Antisocial Behavior (Firesetting) and Parental Behavior (Rejecting, Abusive, Neglectful) and you posit that rejecting behavior of parents toward children is more common in children who set fires, then your hypothesis might be stated as: "Firesetters are more likely to come from families in which parents often exhibit rejecting behavior toward the child." An alternative hypothesis might be that children who are not monitored by their parents (Neglected) are more likely to have the opportunity to develop a behavioral pattern of playing with matches leading to a later behavioral pattern of fire setting. If you introduced mediating constructs, for example, children who set fires are **hostile** because of their experience of parental rejection, hypothesis formulation and testing can become quite complicated. Thus, this model may imply that children who come from rejecting families are more hostile whether or not they set fires. You may also test whether all firesetters are hostile (whether their hostility derives from some source other than parental rejection). **Proposed research that does not explicitly state the hypotheses derived from relationships posited in the conceptual framework that guide the research has a low probability of receiving a high enough priority score to be funded.** Hypotheses should be specific enough that they are disconfirmable and not so general as to be uninteresting. They should also not be so implausible as to lead

reviewers to question their "face validity" (e.g., "(All) Sexually abused children will develop serious mental health problems in adulthood." or "Physical aggression in intimate relationships almost always escalates into battering.")

Moreover, some prioritizing or ranking in terms of importance of hypotheses should be done. Proposals including a long "laundry" list of "hypotheses" between variables (especially when many are obvious or uninteresting) are viewed as negatively by reviewers as are empirical data "fishing" expeditions (i.e., studies that collect a lot of data without any strong theoretical guidance to see how groups differ or how variables are related in the sample). You should highlight the crucial or interesting hypotheses that would have the greatest impact on your research field. Hypotheses that are minimally informative or uninteresting should be backgrounded or not even mentioned.

It is also important that you provide some basis for the validity of the proposed hypotheses. It must not appear that your hypotheses are arbitrary or drawn out of thin air. Rather you must provide some rationale for why the study hypotheses are likely to be confirmed by the research. The types of justification for your hypotheses that you can marshal include (in order of decreasing cogency) (1) previous research you have conducted testing this or similar hypotheses, (2) previous research reported in the scientific literature by other researchers, (3) pilot studies that provide some data on the likely validity of the hypotheses, (4) studies conducted by yourself or others that confirm related hypotheses in your area or analogous hypotheses in other areas, (5) a firm basis for the hypotheses in the conceptual framework for the study (such that the particular hypotheses would largely confirm or disconfirm the particular conceptual framework you are proposing), and (6) your clinical or informal observations.

You should consider whether the research design, sampling plan, measurement plan, and data analysis plan are adequate to actually test the specific hypotheses derived from your conceptual model. There can be many limitations in the research that can compromise your ability to test the hypotheses that you want to test (e.g., selection bias in sampling, use of a cross-sectional design to test posited longitudinal relationships, absent of a necessary comparison group to control for an important confounding variable, or violation of distributional assumptions for a statistical test of the hypothesis). Such limitations should be corrected or acknowledged and discussed in the application.

If you are drawing causal inferences from your data (either probabilistic or deterministic), you should carefully consider the adequacy of such inferences in term of the criteria for making causal inferences (e.g., sufficiency, necessity, temporal contiguity; see, for example, J. Mark Elwood, Causal Relationships in Medicine. Oxford U. Press, 1992 or Clark Glymour, et. al., Discovering Causal Structure. Academic Press). You should entertain alternative explanations for (causal) relationships among variables, especially the existence of "confounding" variables. "Confounding" arises when the relationship between variables can be explained, wholly or partly, by their relationships to other variables. For example, suppose you hypothesize that physical child abuse is "related" to externalizing behavior problems. It may be the case that externalizing behavior problems are more related to other constructs (e.g., constitutional impulsivity, witnessing an

aggressive model, frustration from poor school performance) rather than to the actual experience of physical abuse. Often such potential confounding variables are not measured. Confounding variables are usually controlled either by (a) measuring potential confounding variables and examining their effects relative to the effects of the variables you hypothesize to be related or (b) introducing control or comparison groups into the research to attempt to control confounds (for example, in the above example, comparison groups positive on the confounding variable, but negative on the hypothesized variable, such as nonabused impulsive children, nonabused children witnessing spousal abuse, or nonabused children with poor school performance).

Sometimes confounding variables are introduced into the research because of sampling biases. This is especially likely to happen with volunteer subjects or referred subjects. For example, if one wanted to study the effects of sexual assault on psychiatric symptoms and recruited subjects either from community advertising or referral from a rape crisis center, the inferences one can draw about the relationship between sexual assault and psychiatric symptoms is likely to be distorted to an unknown extent by characteristics of the obtained sample that are not characteristics of all victims of sexual assault. For example, in this case, community volunteers may be less symptomatic than most sexual assault victims because they are willing to participate in a research project that focuses on their victimization or more symptomatic because they are seeking help through the vehicle of research participation. Similarly, the referred sample may be more symptomatic than sexual assault victims in general because they may represent the more severe cases or less symptomatic because they are capable of seeking help. In truth it is difficult, if even possible, to obtain unbiased samples in research involving extant groups and to assess the nature, extent, and effects of sampling biases. Researchers should, however, attempt to minimize sampling biases to the extent possible either by (a) trying to draw the most representative samples possible (e.g., include both referred and nonreferred subjects), (b) by including controls for important biases in the design of the study (e.g., use of control or comparison groups) (c) by measurement of relations between key biasing and study variables (e.g., showing a strong relationship between severity of victimization experience and severity of symptoms in the study sample), or (d) by statistical corrections (e.g., use of stratification or covariance analysis).

It is important to consider and address the temporal relationship among variables that are hypothesized to be causally related (see, e.g., Janice R. Kelly and Joseph B. McGrath, On Time and Method. Sage, 1988). Studies with cross-sectional designs are inherently limited in their ability to establish causal connections among constructs because the hypothesized effective and affected variables are not temporally ordered. Similarly, longitudinal research may be similarly limited if the temporal relationship among variables is problematical (e.g., variables that are temporally distant, such as childhood experiences and adult outcomes; variables hypothesized to have an antecedent and consequence ordering but measured contemporaneously in a longitudinal study; variables assessed retrospectively that are hypothesized as causally related to contemporaneous variables). The limitations in drawing causal inference between constructs due to the temporal relationships among the measured indicator variables should be either

compensated for in the design and analysis of the study or acknowledged as limiting the strength of the causal inferences that can be drawn from the study.

COMMON WEAKNESSES IN THE CONCEPTUAL FRAMEWORK OF PROPOSALS

1. **Conceptual Framework Not Clearly Specified:** The exact set of constructs included in the conceptual framework not clearly identified; individual constructs vague or overly general, not clearly defined or delimited; hypothesized relationships among construct not clearly stated
2. **Too Many Constructs Included In The Conceptual Framework:**
Variables not differentiated into key or important variables and background or less important variables, usually results in inadequate discussion or delineation of the key constructs in the research; variables extensively discussed in literature review that play no role in the current research; inadequate discussion of hypothesized relationships among variables because of the multitude of possible relationships among a large number of variables
3. **Conceptual Framework Overly Simple:** No consideration or discussion of additional variables which should be included in the conceptual framework or which exert a significant impact on the variables already in the conceptual framework; input-output model with no moderating or intervening variables; no differential hypotheses about relationships involving important subcategories of a construct
4. **No Specific Hypotheses About Relationships Among Constructs:** Instead investigator proposes to collect data on a set of variables and describe the empirical relationship among the variables based on the sample data
5. **Too Many Hypotheses:** Especially when many are uninteresting or uninformative
6. **Failure To Consider Alternative Hypotheses To Explain the Results of the Study:** Results due to unmeasured confounding variables, sampling bias, restriction of range of dependent or independent variables; sample size too small to detect true differences or too large to discriminate important from trivial differences
7. **Failure To Discuss Temporal Relationships Among Variables Or To Justify Hypothesized Temporal Relationships:** Failure to distinguish the state versus trait nature of variables; justification for

intervals between data collection points; no attention to or control for temporal contiguity between variables.

Measurement

The measures of the independent and dependent variables used in the research should operationalize the key constructs in the conceptual framework underlying the research and must be fully described; in particular the psychometric adequacy of each measure (reliability and validity) should be reported.

Measurement to Operationalize Constructs

The two major problems seen in the measurement plan of research proposals are (1) failure to include measurement instruments or procedures that adequately measure the key constructs on which the research is based and (2) inclusion of measurement instruments or procedures in the study that measure variables that are not adequately discussed in the application or that measure constructs that are unrelated to constructs in the conceptual framework.

One of the common weaknesses of applications is the failure to specify an adequate measure for an important construct or to provide a justification for the specific measures used in the research in terms of the constructs guiding the research. Sometimes a key construct may be measured by a single question or a few items on a scale or with an instrument or procedure of no known validity or reliability. We also see instances in which a key construct is measured by an instrument that measures a heterogeneous collection of variables rather than by an assessment instrument or procedure that focuses on the specific construct of interest (e.g., use of a very general child problems measure, such as the Child Behavior Checklist, when the major focus of a study is on peer aggressiveness as an important outcome of child physical abuse). Often no indication is given as to why a particular measure was chosen or if it is a good indicator of the construct of interest in the sample.

Also, often available instruments or standardized measurement procedures may have a number of subscales, ratings, or coded categories which are intended either to measure a complete domain, (e.g., all the types of "behavior problems" the measurement developer could think of or factor analyze out) or to measure related constructs. Some or most of the subscales, rating, or coding categories may be irrelevant to the conceptual framework and hypotheses of the research. In which case inclusion of all the subscales (and possibly a "total" score) in data collection can sometimes lead to concerns by a review committee about (a) exactly what measure(s) obtained from the instrument or procedure will be used in the data analysis, (b) a proliferation of variables (especially dependent variables) in the study with attendant sample size and experimentwise error rate problems, and (c) the relationship of the measures to each other and to the conceptual framework and hypotheses of the study. Sometimes it is permissible to omit the superfluous subscales, ratings, or coded categories from the data collection and analysis, but in other cases such omission would have unknown effects on the psychometric properties of the

measurement instrument or procedure. At a minimum you should indicate precisely which measures from multi-measure scales, ratings, and coding schemes will be used in the data analysis.

If there is more than one possible instrument available to measure constructs in your study (e.g., measures of self-esteem or structured psychiatric interviews to determine diagnoses), and, particularly, if there is controversy or disagreement in the field about which among various available measures to use, it is important to anticipate the need to justify the particular measure you propose to use over other possible measures. Such a justification can be developed either from your own knowledge or experience or from consultation with other researchers who have considerable experience with or knowledge of the adequacy of the available, competing measures.

In general, review committees prefer multiple measures of the key constructs involved in the conceptual framework of the study, especially measures from more than one source. The set of measures chosen for the study can be criticized as leading to results based on consistency or bias due to the method (method variance) or source (subject variance) of the measures rather than due to real relationships among the underlying constructs. Thus, use of the multiple forms of the same measurement procedure, such as multiple behavior ratings, or use of multiple measures of a construct obtained from the same data source or reporter, such as various types of self-report measures, can introduce distortion or bias into the relationship among variables due solely to the measurement process. On the other hand, use of multiple independent measures can create difficulties in the analysis of the data (see section later on "multiplicities"), especially when there are discrepancies between informants or low agreement among measures. Thus, when multiple measures and informants are used, one needs to thoughtfully consider how such potential difficulties will be handled. Expert consultation from design, measurement, or data analysis consultants may be helpful in formulating strategies to manage these potential difficulties.

The second common problem seen in the measurement plan of research proposals is the inclusion of a large number of measurement instruments in the research protocol that do not correspond to constructs discussed in the theoretical framework of the study. Usually, these additional measures are included for "exploratory" purposes (i.e., to see if anything turns up in the data by looking at the data from these measures). Such data mining studies are unlikely to receive a high priority score from a review committee because the committee won't know if anything will, in fact, turn up from these unmotivated exploratory analyses and there is a high probability that, with a large number of variables, adventitious results will turn up resulting from chance alone solely due to the large number of measures being examined. There are a few circumstances in which a larger number of variables can be usefully included in the research design and data collection. One of these circumstances, the use of multiple indicators of an underlying construct, has already been mentioned. Another circumstance is when additional information is collected from subjects to aid in the interpretation of results of the study. This might take the form of collecting background information on subjects in naturally occurring groups to determine if the groups differ in important ways; this information can then be used to

temper interpretation of results or incorporated into the analysis in the form of, for example, covariate analysis or post hoc stratification of the groups on an important background variable to determine if group differences are constant across strata. Even with the useful addition of other variables, it is important to keep in mind that the review committee will primarily view the adequacy of the conceptual framework in terms of how well the major, key constructs are defined and delineated and how well the relationship among the key constructs are specified and justified.

Psychometric Adequacy of Measures

There are three ways to obtain measures to operationalize the constructs which play a role in your research: (1) use an instrument or procedure already developed, (2) adapt an existing instrument, or (3) construct your own instrument or procedure.

For any measurement instrument or procedure, whether already developed, adapted, or developed for the research, you will need to establish the reliability of the measurement instrument or procedure. In addition you should consider whether the chosen instruments adequately measure the constructs in your conceptual framework. The basic problem with measurement error introduced into a study by the use of unreliable or invalid instruments is that it reduces the ability of such measures to indicate valid relationships among the constructs in your study. Moreover, differential measurement error in different groups or across the range of a measuring instrument can lead to incorrect conclusion about the real relationship among groups or of the measure with other variables. If you are not knowledgeable about the technical aspects of psychometric properties of measurement instruments, it would be worthwhile to obtain assistance from a psychometric consultant in determining the adequacy of the measures you have selected for your study. An excellent, brief introduction to psychometric issues is Hoi K. Suen, Principles of Test Theories (Erlbaum, 1990); a more advanced treatment of reliability issues is provided in Graham Dunn, Design and Analysis of Reliability Studies (Oxford University Press, 1989).

For many often-used measurement or assessment instruments there are published studies of the psychometric adequacy of the measure (especially of reliability). It is a good practice to search electronic research literature data bases using each of the key measures used in the research as a search term to locate studies that have used these measures or that report psychometric studies of the measures. This literature should be cited in the justification for the choice of a measure for constructs in your studies. It is also a good practice to contact researchers in your own or related fields who might have used measures of the same or similar constructs you are investigating to inquire about their experiences with measures or their opinions about the strengths and weaknesses of the available measures.

Often available measures, especially standardized assessment instruments, are not appropriate measures for the important constructs in the proposed research. In such cases you might propose to adapt an existing instrument by modifying its contents or measurement procedures or develop your own measure. Use of an instrument or

measurement procedure that you have modified, developed, or are developing yourself can be problematic for the review of the application, especially if the instrument measures an important independent or dependent variable, since usually such instruments are of unknown validity and reliability. In such cases it is wise to specify in considerable detail the steps that will be taken in the research to develop and evaluate the proposed instrument; this is especially true when you proposed to use a questionnaire or (semi-) structured interview to collect important information. It is preferable to develop a preliminary version of the instrument and at least collect some pilot data on the reliability and validity of the new or modified instrument. For brief introductions to procedures used to develop assessment instruments, see René Dawis, Scale construction. Journal of Counseling Psychology, 1987, 34, 481-489 or John Rust and Susan Golombok (1989), Modern Psychometrics (London: Routledge). For a model of the psychometric process of developing a report measure see Angold, A., et.al. (1995) Development of a Short Questionnaire for use in epidemiological studies of depression in children and adolescents. International Journal of Methods in Psychiatric Research, 5, 237-249, and Messer, S.C., et.al. (1995) Development of a Short Questionnaire for use in epidemiological studies of depression in children and adolescents: Factor composition and structure across development. International Journal of Methods in Psychiatric Research, 5, 251-262.

It is often appropriate or necessary to conduct small pilot studies of the psychometric adequacy of measures you intend to use in your research. This is particularly the case when (1) you modify a standard instrument to increase its face validity or to get information more relevant to your research questions (e.g., change wording of items in a structured interview), (2) you modify procedures in order to use the instrument with a different population or in different circumstances than were used in the original development of the instrument (e.g., using pictorial representation of verbal items with young children or poor readers), (3) you extend use of an instrument to groups that either were not in the norming group on which the psychometric adequacy of the instrument was assessed or who were under-represented in the norming group (e.g., a younger or older age group or an ethnic minority group), and (3) you construct your own assessment instrument for one or more key constructs in your research (e.g., a questionnaire or semi-structured interview). Such pilot studies do not require a large number of subjects and may possibly impress a review committee concerning your attention to this important matter. (see Chapter 1 in Joseph L. Fleiss, The Design and Analysis of Clinical Experiments, Wiley, 1986, for an introduction to such pilot studies of measure reliability.)

You should carefully consider the implications of published or pilot studies of the psychometric adequacy of the measures you plan to use in your research. Assessment of reliability of a measure can take several different forms depending on the type of measurement obtained and the procedure by which the measure is obtained. Thus, for items making up a (continuous) scale some indication of adequate scale reliability should be presented, (e.g., item-total correlations or Cronbach's alpha); for classification procedures some indicators of the reliability of classifications should be presented (e.g., sensitivity and specificity); and for coding procedures some indication of inter-rater

reliability should be presented (e.g., percent agreement or Cohen's kappa). It is not enough to simply report that there have been studies of the psychometric adequacy of a measure; the results of such studies should be used to evaluate the adequacy of proposed measures for your study. For example, if a proposed key measure in the form of a multi-item scale has been reported to have an internal consistency of around .60 or lower in samples similar to the one you are using in your research, a review committee is likely to (rightly) observe that the particular scale is not entirely adequate as a measure of a proposed underlying unidimensional construct and either the instrument has to be improved or a better measure selected (even if you happen to like the particular scale you propose to use or everyone else uses it).

An important issue in the validity of measurement of constructs is the quality of the data that is available from various sources, especially potential sources of informant bias. Thus, certain informants may have agendas that bias their responses (e.g., abusing parents who are at risk for having their children removed or antisocial individuals may want to portray their children or themselves as less problematic than they are, contrarily, individuals seeking treatment may want to portray themselves as worse off than they really are), or certain informants may have limitations in terms of exposure, mental state, or observational ability interfering with their ability to make informed judgements or to report reliable observations (e.g., foster parents or teachers as informants may have limited experience with a particular child or individual with severe psychopathology may be poor informants about themselves or others), or collection of multiple measures from the same informant is liable to introduce dependencies into the data arising solely from using the same source for these data (i.e., the measures will not be independent, will have correlated errors). Careful consideration should be given to potential sources of respondent bias and strategies to reduce these biases incorporated into the research or the limitations arising from this source should be acknowledge in the discussion.

In addition to potential problems with measurement validity arising from data sources, there are also potential validity problems with measurement instruments or procedures themselves. Because there is much less attention paid to validity than to reliability in psychological assessment, there are often few or no studies available on the validity of measurement scales. Such studies could include factor analyses or item response analyses of potentially multidimensional constructs, correlations of a scale assessing a construct with other recognized scales assessing the same construct, correlations of scale scores with other indicators of the underlying construct, or scale scores differentiating criterion groups differing on the underlying construct. You should be aware of the validity literature on measures you use in your study, especially for commonly used measures in a particular research field, because this sometimes presents review problems for a proposal (e.g., when a knowledgeable reviewer is aware of studies that indicate that a measure you propose to use does a poor job of differentiating criterion groups related to the underlying construct).

Absent a literature on the validity of a particular key instrument or measurement procedure, reviewers will sometimes assess the **face or content validity** of the assessment instrument. They might form an opinion about whether the content and

number of items in your assessment instrument or procedure is likely to adequately assess the underlying construct. You should scrutinize your proposed instruments from this same perspective. For example, some researchers have criticized the widely used original Conflict Tactics Scale, which assesses verbal and physical aggression in relationships, in terms of the adequacy of the content of the items for certain types of research (i.e., the items don't differentiate defensive or retaliatory aggression from offensive aggression, don't adequately assess the severity of injury or potential injury arising from physical aggression, and don't include items for some more extreme forms of physically abusive behavior (e.g., scalding or purposely injuring pets)). The revised Conflict Tactic Scale was developed to address some of these measurement concerns.

Additionally, attention should be paid to **demand characteristics** of items or measurement procedures, especially involving socially desirable responding by informants or a press for consistency in responding (e.g., parents seeking treatment for their child are unlikely to rate their child as not having any problems on a behavior problem checklist).

You should consider the adequacy of the measures for the particular populations you are investigating (e.g. a measure with psychometric adequacy on a normative sample might not be adequate or of unknown adequacy with, for example, younger or older subject, subjects of different ethnicity, or members of special populations, such as mental patients). You should also consider the appropriateness of the measurement with regard to the intelligence, reading ability, ethnicity, and social class of the subjects.

In research involving assessment of children and adolescents, developmental issues in assessment should be carefully considered. The review committee is liable to raise questions about the measurement and interpretation of responses given by children from a wide age range (e.g., 6 to 17). Often researchers will use such a broad age range because they do not have access to a sufficient number of subjects in a more restricted age range. This limitation in subject recruitment can create problems for the research because it raises a host of issues - conceptual (e.g., developmental differences are often not addressed), measurement (e.g., often norms for an instrument are available for only part of the age range), and data interpretation (e.g., combining results over a broad age range may average very different response patterns at different ages). There are additional measurement issues related to different ages of subjects (e.g., in general, measures are less reliable the younger the age of the subject and, contrarily, parent-child agreement on many measures declines with age) that should be considered and, possibly, discussed in the application if they are likely to have a significant impact on the research.

Studies that propose to use "official" records should provide some documentation of their access to such records and an indication of the quality of information that can be obtained from such records. With questionnaires, (semi-)structured interviews, and coding schemes, procedures to establish inter-rater (or interviewer) reliability should be given.

COMMON WEAKNESSES IN THE MEASUREMENT PLAN OF PROPOSALS

1. Measures Not Linked To Constructs In the Conceptual Framework:

Measures are used that measure constructs not discussed in the conceptual framework, particularly likely with commonly used measures (e.g., SCL-90 or CBCL); collection and use of all measures in a battery of measures when only one measure in the battery corresponds to a construct discussed in the conceptual framework

2. Key Constructs Not Adequately Measured: Omission of a measure for key constructs discussed in the conceptual framework or use of a measure that has poor psychometric properties (e.g., low reliability) or that does not measure the underlying construct discussed in the conceptual framework (e.g. use of an achievement test to measure "intelligence") or use of a "face valid" measure developed by the researcher or others with no reliability or validity studies of the measure's adequacy

3. Failure To Provide Psychometric Data For Measures: No citation of psychometric studies of measures; assertions that measures are psychometrically adequate without citing specific data or reports as justification for the claim; omission of reports contra-indicating use of a measure for a construct or in certain situations or with certain populations

4. Failure to Consider the Psychometric Adequacy of Measures For The Specific Proposed Study: Citation of psychometric data collected on general samples that might not be applicable to the special populations (e.g., children under a certain age, people with limited reading ability, institutionalized mental patients) or circumstances (e.g., acute phase of a disaster or trauma, in an institution, or abusive family) characterizing the proposed research

5. Uncontrolled Measurement Biases: Use of individuals with "vested interest" as data collectors (e.g., therapists as raters of treatment outcome, researchers as raters of group differences); shared variance in measures arising from sources other than the relationship among the measured constructs (e.g., variance due to multiple ratings from the same raters, measures with different names measuring the same underlying construct); lack of blindness of data collectors or providers to the research hypotheses or group membership of study participants.

Sampling plan of the study

The sampling plan of a research study addresses two major issues: (1) the *representativeness* of the research sample to a theoretical population of interest and (2) the *generalizability* of the research results (i.e., the population range to which research results would be applicable). In general, research that claims to describe only a particular (limited) sample is not of as much scientific interest as research that attempts to infer

characteristics or relationships in more general populations. Target populations are of two general types: area populations, defined by geographical area of residence or location (e.g., the U.S. population, population of a particular community), and special populations, defined by some characteristic of population members that is not necessarily geographically located (e.g., subjects in a certain age range, populations exposed to or exhibiting a particular disease or having experienced or been exposed to a particular type of risk factor). Clinical epidemiology attempts to estimate the incidence and prevalence of "diseases" or clinical conditions and to understand the etiological, risk and protective factors for the onset, course, remission, recurrence and outcome of diseases or clinical conditions. Mental health epidemiological research conducts studies of (1) mental health problems, disorders, or symptoms in general community, areal, or national populations and/or of risk or protective factors for such mental health problems and of (2) etiology, characteristics and course, risk and protective factors, and outcomes of mental health problems in clinical populations defined by such characteristics as having a diagnosed mental disorder or range of mental health problems or symptoms, evidencing or exposed to risk factors associated with mental health problems, or seeking treatment for mental health problems. Definitional issues and sampling strategies differ for these two types of populations.

In both types of studies you should give some thought to how general is your study population. Research based on limited sampling is less likely to be highly evaluated by a review committee. In general review committees prefer samples representative of larger, more general populations than of more restricted populations. Thus, nationally representative samples, community samples, multi-site clinical samples, and single-site clinic or institution samples represent declining representativeness and generalizability. Often, however, samples representative of larger, general populations are unsuitable when the intent of the study is to study low-base rate behaviors (e.g. deviance or severe antisocial behavior) or clinically significant conditions (e.g., studies of diagnosed mental disorders). In such cases, general population samples are unlikely to yield sufficient numbers of subjects exhibiting the low-base rate behaviors or disorders or will include mainly subjects exhibiting the less severe forms of the low-base rate behaviors or subclinical manifestation of the disorder. Thus, it may be necessary to adopt sample recruitment procedures to increase the number of individuals exhibiting the low-base rate behavior or clinical condition (e.g., recruitment at a facility seeing or providing services to such individuals, such as a rape crisis center for sexual assault victims, or screening a more general population, such as screening high school students for antisocial behavior). Even in such cases it is important to consider and discuss sampling limitations in the proposal.

Proposals have received poor priority scores solely due to problems of unrepresentativeness or generalizability. For example, the generalizability of study of antisocial youth in which the sample is drawn from a high income area is likely to be questioned due to the unrepresentativeness of these youth compared to antisocial youth in general. Similarly, if you propose to study sexual assault among women and you recruit a representative sample of all women enrolled at a particular college or university; you can not validity claim that the results of this study will generalize to all women who have

been sexually assaulted (e.g., women who don't attend college or in a different age range than predominate in college populations). Such a study may be valuable in its own right (e.g., as a study of a population at a high-risk for sexual assault), but some consideration should be given to the limits of generalizability of the sample and discussed in the proposal. This is particularly true when your conceptual framework and research design seem to imply that the research will address a general population (e.g., rape victims or antisocial youth), but the available subjects are an obviously biased sample from the larger general population. For example, the above two examples would not generalize to sexual assault victims who do not seek treatment nor to antisocial youth who are not in school, respectively.

There are some circumstances in which non-representative or non-generalizable study samples are acceptable. In general population studies a census or near-census of a relatively rare population or a model-testing or cross-validation study of a large, but biased, sample, in which, other than sampling, the research is particularly strong in conceptualization and methodology, are examples of studies which might be favorably reviewed despite sampling limitations. In many clinical studies samples are recruited such that groups that are compared differ on some clinical characteristic (e.g., exposure to a risk factor) and are matched on other characteristics such that the goal of the study is to examine the effects of the clinical characteristic given other factors are comparable. In such studies sampling is less of an issue, but such studies can still be criticized for involving non-representative or non-generalizable sampling (e.g., comparing depressed patients with matched non-depressed patients drawn from practices of private psychiatrist would not necessarily be generalizable to patients who receive care in other settings or who are not receiving treatment).

Sampling designs. A sampling design is a procedure or set of procedures that identifies subjects eligible for participation in the research. The sampling plan of the study draws participants from several different types of populations of decreasing scope. The *target population* represents the population that is the focus of the research and defined in the conceptual framework of the study.

The *sampling frame* of a study is the set of subjects from the target population that are eligible for participation in the study due to the sampling methodology of the study. In area population studies the sampling frame is determined by the methods by which potential subjects are contacted (e.g., household surveys in larger sampling units, such as blocks, neighborhoods, or census tracts, random, or stratified sampling of schools to identify children for research studies, or random digit dialing for telephone interviews). In special population studies, potential subjects are identified either by screening in area populations or by recruitment from data bases or institutions that identify, track, or provide services to individuals in the special population. In general, the sampling frame of the study restricts, to a greater or lesser degree, the target population to a smaller population that the study has access to with concomitant restriction on representativeness and generalizability. Thus, it is quite common for clinical studies to recruit subjects who have been diagnosed with or are seeking treatment for a disorder or problem from a clinical facility. Such a sampling source does not access individuals with the disorder or

condition who have not been diagnosed or do not seek treatment and who might be quite different (e.g., in severity of symptoms or functional impairment) from diagnosed or treatment seeking individuals. Similarly, if the theoretical framework in a study focuses on "abused children" as the target population, but, because of access problems, only children who have been reported to a protective service agency and confirmed for abuse are eligible for inclusion in the study, then the sampling frame of the study would include only "abused children reported to and confirmed by a protective service agency." The sampling frame in this instance would exclude non-reported abused children and reported, abused children who were not confirmed. Another example would be a study of an area sample by random digit dialing of telephone exchanges in a certain geographical area. Such telephone surveys exclude individuals who do not have a telephone from the sample frame. In developing a grant application, the applicant should be mindful of the restrictions in representativeness and generalizeability with respect to the (conceptual) target population introduced by the sampling frame of the study. If possible, these restrictions should be analyzed, assessed, and minimized and their effects on the representativeness and generalizeability of the research should be acknowledged as a limitation of the research.

The *study sample* is a subset of the sampling frame that is actually included in the research study. The study sample would not include members of the study frame who refuse to participate in the research, cannot be contacted, or are not eligible for participation for other reasons (e.g., non-English speakers). In general the target population of the study is larger than the sampling frame of the study which is larger than the study sample. As discussed above with respect to the sampling frame of the study, eligibility restrictions and refusals to participate in the research by eligible subject can seriously restrict the representativeness and generalizeability of the study.

Eligibility inclusion and exclusion criteria should be motivated by the conceptual framework of the study or by data collection or procedural issues and clearly stated and operationalized. This is particularly important in clinical studies in which there are a number of issues relevant to the specification of inclusion and exclusion criteria for "cases." Issues of inclusion criteria include (1) how cases are identified and (2) the role of symptom severity and clinical significance. Cases may be identified as meeting diagnostic criteria for mental disorders, evidencing symptoms, experiencing problems in functioning, or seeking treatment. In such cases the specific procedures by which "cases" are identified should be detailed. For example, individuals may meet diagnostic criteria either by consensus or blind agreement of experienced clinicians, a clinical judgement of a single clinician, classification algorithms of a structured or semi-structured interview (e.g., a DIS or SCID interview), or by being above a cutoff score on a symptom inventory. Inclusion criteria may be determined by theoretical or clinical interest. Thus, some studies may include only individuals who make diagnostic criteria; other studies also require that cases meet criteria of functional impairment or clinical severity. Similarly, some studies may require cases to meet diagnostic criteria currently; others may include individuals who have met diagnostic criteria in the past or who evidence subsyndromal symptomatology. Some issues of exclusion in clinical studies are (1) whether to exclude individuals who evidence limitations that may interfere with data

collection (e.g., acutely psychotic individuals who may not be able to provide accurate data, individuals exposed to certain types of drugs or other circumstances that may interfere with collection of biological response data, or individuals with intellectual, language or readings problems or difficulties); (2) whether to exclude individuals with co-morbid conditions as sometimes interest is in the effects of the specific condition uncomplicated by other detrimental conditions or co-morbid conditions may complicate or compromise data collection or interpretation of results (e.g., to include individuals with co-morbid substance abuse in a treatment study for depression); and (3) whether to include treated individuals in the sample as, presumably, sometimes treatment may alleviate or control the effects of psychopathology that are of interest (e.g., antipsychotic medication may improve deficient attentional processes in schizophrenics). As with inclusion criteria, exclusion criteria must be explicitly operationalized. Thus, a researcher who excludes "psychotic" individuals in a study of mental patients should specify the criteria that will be used to determine psychosis and how they will be applied to determine if a patient is "psychotic." Some researchers impose so many exclusion criteria on their sample subjects for data collection reasons that the representativeness of the sample is seriously affected. For example, clinical researchers might exclude individuals from a study with co-morbid disorders when the clinical population being studied has high rates of co-morbidity (e.g., combat veterans with co-morbid PTSD and Major Depressive Disorder) or might exclude individuals from the clinical population receiving certain types of commonly prescribed medications or having certain types of biological statuses (e.g., alcoholics) that might interfere with biological assessment procedures or might exclude individuals receiving treatment outside the research study.

Every effort should be made to minimize participation refusal by eligible subjects. Low percentages of potential subjects who agree to participate in the study or differential rates of consent to participate in different study groups can seriously affect the representativeness of your samples vis-a-vis the populations they represent and this issue should be addressed rather than ignored. This is particularly true of hard-to-recruit populations, such as child and spouse abusers, sex offenders and other criminals, or mental patients, who might be suspicious of requests for participation in "research". Studies with high rates of refusals by eligible participants are regarded as seriously compromised in their ability to generalize to the target population. Subject recruitment procedures should be adopted that will minimize refusal rates, such as use of service providers to recruit subjects (provided incentive are offered to the service providers so that they are motivated to be cooperative with subject recruitment goals and such recruitment is not viewed as coercive by clients), clear description of the research goals and procedures of the study as part of the recruitment process, collaboration with affected communities and consumers in the development and implementation of the research plan and in the participant recruitment plan, minimizing the participation and response burden on potential participants (e.g., provision of transportation or child care for parents) and payment of reasonable financial incentives for participation.

Sampling procedures. There are two general approaches to sampling from a population: (1) probability sampling and (2) non-probability sampling. In probability sampling every member of the population has a known or estimable probability of being selected in a

sample; in non-probability sampling the probability of selection is unknown for most members of the population. There are several types of both probability and non-probability sampling.

For probability sampling, if there is a list of population members (e.g., all fire burn victims in a city in a year period), they can all have an equal probability of being (randomly) selected by means of *simple random sampling*. Sometimes population units may be selected with known but unequal probabilities (e.g., where selection probability is related to some general measure of size, such as if patients in a medical study are selected with a probability related to the number of medical visits or length of hospital stay). If a list of population members does not exist, survey researchers will often use other means of selecting subjects and assigning selection probabilities to potential participants. Two of the most common procedures are *cluster sampling* and *stratification*. In cluster sampling, a list is not available of all population members but there is available a list or other means of exhaustively specifying larger population units (e.g., usually geographically located clusters such as households, blocks, or schools). Clusters are first sampled then individuals are sampled from clusters. In stratification, the distribution of some defined characteristics of the sample are known for the general population (e.g., ethnicity or age in general population studies) and participants are screened or otherwise selected by means of the strata that characterizes them. More complex sampling plans may combine sequentially different types of sampling (e.g., stratification, then clustering, then screening, then simple random sampling) in *multistage sampling plans*.

In probability sampling the probability of selection can be ascertained or estimated for members of the population. In non-probability sampling, the probability of selection of population members can not be estimated. **Thus, with non-probability sampling, the sample in the research study does not have a known probability of representing any larger population and non-probability samples cannot be claimed to probabilistically represent larger populations as is true of probability samples.** Non-probabilistic sampling can lead to samples that are not representative of the theoretical population of interest and produce research results that do not hold in the theoretical population or are otherwise misleading (see discussion of "bias" below).

The question of the sampling representativeness of a sample should be particularly addressed for "convenience" samples (i.e. samples available at a particular clinic or in a school system chosen solely because of investigator access). Review committees are becoming more and more reluctant to award high priority scores to studies involving convenience samples and other types of non-probabilistic samples, such as volunteer or nominated samples. Volunteer or self-selected samples (e.g, subjects who respond to newspaper or other advertisements) are considered to be biased to some unknown extent by unique characteristics of individuals who volunteer for the research as compared to a random sample of individuals from the target population. Similar concerns about sample bias might be raised about "nominated" samples (e.g., subjects nominated to be in a study by therapists or teachers). In general, review committees prefer sample recruitment strategies that allow all individuals in the targeted population an equal opportunity to participate in the research (e.g., locating subject in a community sample who meet criteria

for participating in the study by means of random digit dialing or through targeted contact of randomly selected households in census divisions, recruiting consecutive admissions to a hospital or other institution or program, or identifying potential subjects randomly from an institutional or agency data base). Such restrictions on recruitment strategies can pose a hardship on the researcher, particularly, in studies addressing low base rate behaviors or some types of private behaviors or events not often known by others and may adversely affect the feasibility of the study.

Some types of sampling plans that can argue for representiveness of the research sample include: sampling that attempts to recruit typical cases, sampling that screens and selects cases on the basis of theoretically important stratification variables (e.g., ethnicity or severity of problem), sampling plans that assume a model of the distribution of population values or of the relationship between population values and obtainable auxiliary population information and recruits a sample that matches the posited population distributions, and quota sampling to identify groups of participants who meet particular important or common profiles of population members (e.g., homeless single women with children, homeless families with children, homeless single adults with a history of mental illness). Even in such non-representative samples it is important to attempt to reduce potential sources of bias by sampling probabilistically from eligible members of the sampling frame. Thus, including consecutive admissions or every nth admission to an inpatient unit is likely to lead to a less biased sample than recruiting participants through nominations by admissions workers or clinicians.

Non-probability sample are often the norm in studies of hard-to-identify or hard-to-located populations (e.g., "abused children," "battered women," or "homeless families"). If you cannot acquire a probability sample of the particular hard-to-access population on which you wish to conduct research, you should acknowledge the probabilistic non-representativeness of the sample and provide an argument for either representative of the sample on other grounds or the applicability of the results from your sample to the general population or to other important samples. At a minimum you should obtain data describing the characteristics of the samples that you will have access to (e.g., demographic characteristics of families or children, referral reason for clinic samples, and diagnostic information on psychiatric inpatients), so that these characteristics can be compared to the population or a typical sample of interest. The available sample should be compared with the target population on available important population characteristics from other studies. A discussion of procedures to increase or ensure the representativeness of the sample vis-a-vis the target population (e.g. a multi-site study) would be helpful. Often the representative of the sample can be increased by use of sampling procedures that have been developed to sample rare, difficult-to-detect, and other types of special populations. Such sampling procedures include use of multiple lists and frames, specialized screening procedures, use of partial lists and screening populations for non-listed members, cluster sampling techniques for concentrated populations, network and snowball sampling procedures, site sampling, and combinations of these methods (for an introduction, see Thompson, S.K. (1992) Sampling. New York: Wiley.; Sudman, S., Sirken, M.G., and Cowan, C.D. (1988) Sampling rare and elusive populations. Science, 240, 991-996; Sudman, S. and Kalton, G. (1986) New

developments in the sampling of special populations. Annual Review of Sociology, 12, 401-429; Kalton, G. (1993) Sampling considerations in research on HIV risk and illness. In D.G. Ostrow and R.C. Kessler (Eds.) Methodological Issues in AIDS Behavioral Research, New York: Plenum).

Recruitment procedures. How subjects are actually recruited into the study operationalizes the sampling design of the study. The recruitment procedures that will be used in the study should be spelled out. How and who will identify potential subjects, how potential subjects will be approached and recruited, and the criteria for inclusion and exclusion of potential subjects and the rationale for these criteria should all be spelled out.

To demonstrate the feasibility of your project you should provide documentation that you can, in fact, obtain the numbers of subjects you specify in your various study groups. This documentation usually takes the form of figures on the number of subjects available in the total subject pool and in pools for various subgroups, including figures or estimates from previous research of the percentage of subjects in the various subject pools who are likely to choose to participate in the research. For example, if you propose to include 50 male abuse victims in a study group, the agency with which you have cooperation might have 200 abuse victims, but only 20 males who meet study eligibility criteria available in the study period. As part of your documentation, you should also include letters of support or cooperation from agencies which will provide you with access to subjects or to data.

Specific, operational criteria for selecting and excluding subjects should be stated. For example, if your study involves antisocial youth or sex abuse victims, what criteria will you use to choose these subjects? The definition and feasibility of using these criteria can be pivotal in the review of your proposal. For example, if you propose to study "serious antisocial youth" and your inclusion criterion is "ever been arrested for a delinquent act" the review committee is likely to object to the adequacy of this criterion to define "serious antisocial youth", since many such individuals have never been arrested and most juvenile arrestees are arrested for less serious offenses. Similarly, , if you propose to obtain your sample of "serious antisocial youth" from middle or upper income communities, the committee might object that you are unlikely to find a sufficient number of "serious antisocial youth" in such low risk communities.

Particular attention should be paid to false positives and false negatives in study and control groups. For example, subjects selected as "antisocial youth" selected from a pool of juvenile arrestees with a criteria of "committed a serious delinquent act" would be false positives if, in fact, they were innocent of committing the alleged offense. Conversely, subjects selected as a "non-abused control" would be false negatives, if they had been abused, but never reported. Exclusion criteria might also include important confounding variables which could affect results of the study, (e.g. low IQ in a study of school achievement of abused children or out-of-home placement in a study of families of delinquents).

Sampling considerations in data analysis. Epidemiological studies attempt to estimate population characteristics from sample data. Estimates of population characteristics (e.g., population totals or means on some variable of interest) can vary with the sampling probability of population units (e.g., simple random sampling with or without replacement vs. unequal probability sampling), with the sampling design (e.g., simple random sampling vs. cluster sampling) and with the use of additional population information (design-unbiased probability sampling estimators vs. model-based estimators, such as ratio and regression estimators). Estimators have two properties that determine their relative merit: *bias* and *precision*. An unbiased estimator has the property that the expected value of the estimator equals the population value over all possible samples (alternatively converges to the population value as sample size increases). Bias represents the difference between the expected values of the estimator and the population value. The precision of an estimator refers to the size of the standard error or the width of the confidence interval of the estimator. Good estimators are unbiased with small standard errors, but there exist biased estimators with small standard errors and unbiased estimators with large standard errors. There are two factors that affect the biasedness and precision of estimators: (a) sampling error and (b) non-sampling errors.

Sampling error results from the inevitable loss of precision of using a sample to estimate a population value. In ideal cases of simple random sampling, sample estimates are, in general, unbiased estimators and their precision depends largely on sample size. More complex sampling schemes may be adopted either out of necessity or to attempt to increase the precision of simple random sample estimators. Often more complex estimators will introduce bias into estimates of population values and their value will depend on the relative advantage of increasing precision of estimation versus introducing bias into estimation.

Non-sampling errors are of 3 major types: *frame errors*, including (undercoverage (failure to include eligible participants in sampling frame), overcoverage (inclusion of ineligible in sampling frame), and multiplicities (multiple chances for potential participants to be included in sampling frame)); *nonresponse errors*, including population units that are eligible but unavailable (e.g., population members that can't be contacted or who are non-English speakers) or that refuse to participate); and *measurement errors*, including misclassification or other data reporting or recording errors. Non-sampling errors usually lead to biased estimates of population values and may decrease the precision of population estimates. Remedies for non-sampling errors include (1) preventive measures taken during development of the sampling procedures or data collection to reduce non-sampling errors, (e.g., reducing the percentage of nonidentified, ineligible, and nonrespondents by means of canvassing, screening, and incentives for participation, staff training in and monitoring of recruitment and data collection, and followup efforts to recruit unavailable or refusing eligible members of the sampling frame); (2) attempt to assess and then reduce sources and extent of bias (e.g., put extra effort in recruiting hard-to-locate and hard-to-recruit population members, use additional procedures to recruit a sample of non-identified or nonrespondents and compare the extent of bias, compare sample characteristics to larger samples already collected or to known characteristics of the population, or compare measurement estimates in subgroups

or strata of the sample which might be hypothesized to differ in recruitment probability) and (3) adjust for bias in estimation procedure, (e.g., by determining a scaling factor that varies linearly with recruitment or nonresponse probability and regressing the target measure on the scaling factor and then extrapolating the regression to the nonidentified or nonrespondent population members; two primary methods are *weighting class* adjustment and *post-stratification adjustment*). Remedial procedures for nonsampling errors are discussed in Lessler, J.T. and Kalsbeek, W.D. (1992) Nonsampling Error in Surveys. New York: Wiley.

Sample attrition over time can affect the biasedness of sample estimates of population estimates. The issue of sample attrition should be addressed, particularly, in longitudinal or intervention research and, if it is likely to differ for the different research groups (e.g. a clinical vs. a control sample). Selective or differential attrition opens the possibility that research results (or non-results) are due to resulting biased sample differences in the groups rather than due to true differences between the underlying populations groups or to the intervention (even if the groups were initially representative of the target populations or the intervention groups were randomly assigned). Realistic estimates of attrition together with an explanation of how the estimates were derived should be included. If possible, procedures for characterizing refusers and dropouts and comparing them to sample subjects should be proposed. Longitudinal studies, especially that collect data over a period of a year or longer, should include discussion of tracking procedures to reduce sample attrition over time, since this is one of the most prevalent problems in this type of research. You should give careful thought to the timing and amount of subject payments as an incentive to participation in the study, especially for high risk populations

Inclusion of Gender and Minority Groups

Review committees are now required to examine the participation of both genders and of minorities in the sampling plan for most research studies. If one gender and/or minorities are excluded or under-represented in your planned sample(s) for any reason (for example, the demographic area from which you are obtaining subjects or the gender distribution of the symptom or diagnosis you are studying), you must provide an explicit rationale for the exclusion or under-representation of these groups specifically indicating why this sampling is justified or not remediable. For further information, see the policy notices now included in Public Health Service grant application packets. In addition, prevention and treatment studies are now required to include sufficient numbers of males and females and minorities such that the effectiveness of the intervention for these different groups can be assessed. This requirement is waived only in cases where the disorder or problem does not occur in one or more of the gender and/or minority groups or occurs so rarely that recruitment of a sufficient number of subjects in that group would be infeasible (e.g., women Vietnam veterans exposed to direct combat). In research other than intervention research, review committees expect recruitment of a sufficient number of males and females and minorities to allow at least preliminary comparisons of group differences based on gender and ethnicity, an explicit discussion of expectations about such differences, and specification of data analytic procedures to analyze gender or ethnic group differences in the study. This extension of the basic requirement of ensuring

adequate representation of women and minorities in the study population can have profound effects on the feasibility of the research, especially with regard to sample size when these subgroups are analyzed separately. This issue should be carefully considered.

Inclusion of Children

It is now the policy of NIH that children, defined as under the age of 21, must be included in all human subjects research, conducted or supported by NIH, unless there are scientific and ethical reasons to exclude children and/or adolescents (NIH Guide for Grants and Contracts, March 6, 1998). The applicant should include a section in the application titled "Participation of Children," which should provide either a description of the plans to recruit children and/or adolescents or a rationale for excluding children or a specific age range of children and adolescents. If children are included in the study, there should be sufficient numbers so that study results with respect to children can be analyzed. The NIH Guide announcement list 7 types of rationales from excluding children from a study on scientific (e.g., study hypotheses are not applicable to children or disorder does not occur or occurs very rarely in children) or ethical grounds (e.g., greater than minimal level of risk for participation). Scientific review groups at the NIH will evaluate each application for appropriateness of inclusion/exclusion of children as research subjects. Applications which are judged as "unacceptable" in regard to the age-appropriate inclusion or exclusion of children in the research project will not be funded unless the exclusion of children is remedied or a better rationale provided for the exclusion of children. Additional information on this policy is provided at an informational Website for this policy at <http://www.nih.gov/grants/funding/children/children.htm>.

COMMON WEAKNESSES IN THE SAMPLING PLAN OF PROPOSALS

1. Failure To Adequately Describe The Sampling Plan: Failure to specify how subjects will be recruited; no discussion of inclusion and exclusion criteria for participation; no letters indicating cooperation from participant recruitment sources; no discussion of number of subjects available, especially of various important subgroups in the design
2. Use of Convenience Sample: No justification for use of a convenience sample other than easy access for subject recruitment; use of more than one convenience site with very different subject populations
3. Failure To Address Sampling Bias Issues: No discussion of adequacy of sample as representative of a target population; no description of characteristics of individuals from whom the study sample will be drawn (e.g., characteristics of clients at a clinic providing subjects); lack of procedures to attempt to correct for inherent sampling biases, such as sampling procedures, research design methods (e.g., comparison groups) or statistical procedures (e.g., covariance analysis); no discussion of potential impact of sample biases on results and interpretation of study and

generalizeability of results to other samples; no attention to biases introduced by refusals to participate in the research; no attention to sample attrition, especially selective rather than random attrition, in longitudinal and intervention research

4. No Justification For Sample Size: No discussion of issues of requirements for adequate statistical power; no discussion of sample sizes needed in consideration of precision of measurement of key constructs; no attention to number of subjects in subgroups even though sample size for entire study is adequate

Procedures for Data Collection

The data collection plan or procedures should be specified. This should include an indication of which measures are going to be collected from which subjects on which occasions, the order of administration of the measures, and the time required of each subject on each assessment occasion. An application that is unclear or confuses reviewers about exactly how data is to be collected often fares poorly in a review. If data is being collected at more than one point in time, a justification is needed for the particular time points chosen and the intervals between data collection points, e.g. evidence from the literature on the amount of time symptoms typically take to develop or remit in a trauma group. Similarly, if data is not being collected from all subjects by the same procedures or schedule, a justification for the differences in data collection procedures for different groups of subject is needed.

It is often helpful to construct a chart that clearly shows in terms of a project time line when each task in the project will be initiated, how long it will take to complete, and when it will be completed. Such a chart can also aid you in determining number of personnel and amount of resources needed during each budget period. Such a chart can also help you determine if the project is feasible to accomplish in the time period proposed and with the budgeted staff and resources. If the timeline or number of tasks to be accomplished are perceived by reviewers as being unrealistic, they will tend to view the applicant as unlikely to be able to actually complete the research project as proposed and this judgement will be reflected in the priority score they give to the project. Reviewers often term such proposed research as "overly ambitious." You should especially consider whether the estimates of rate of subject recruitment, time period to completely assess all subjects in the study, time to locate previously recruited subjects in longitudinal research, and

time to accomplish all proposed data analytic tasks are realistic. Often a small pilot study in which subjects are recruited and assessed with the full research battery can both aid in the estimation of the time and resources needed to complete the research and to convince reviewers that the research can be accomplished in the time period and with the resources proposed.

In general you should include a copy of the instruments that are not well-known or that you have developed yourself in the Appendices. Reviewers are likely to look at the individual items in these measures and form an impression of the face validity of the instrument to measure constructs specified in your conceptual framework, e.g., if "previous victimization" is an important construct in your conceptual framework and you use a general "life stress" instrument which contains only 2 general questions about victimization experiences, a reviewer is likely to question the adequacy of this instrument as a measure of the victimization construct. A reviewer may also form an impression of the suitability of the items for particular groups of subjects in your research, e.g. self-report instruments with very young children, reading level of questionnaires for low SES respondents, and family life or "life style" questionnaires with different cultural groups.

Some issues in data collection procedures that may be raised by a Review Committee are:

(1) Familiarity with settings and subjects from which data will be collected and/or the types of data that will be collected. If you have not had experience with collecting data in specialized settings or with specific types of subjects and, especially, if you fail to address the significant difficulties in obtaining quality data in these settings or with these subjects, a Review Committee may raise issue with the feasibility of the study or the suitability of the data for analysis or inference. Some of these difficulties might include obtaining voluntary consent from a significant proportion of available subjects, e.g. psychiatric hospital patients, penal inmates, and abusing parents; obtaining valid and reliable responses, especially self-report data from young children, trauma victims, psychiatric hospital patients, and antisocial individuals; and extracting comparable quality data for different individuals from official records, such as rap sheets or hospital records from different systems. For example, if you propose to collect a large number of self-report measures from psychiatric hospital inpatients and don't address such data collection issues as access to patients, flow of eligible patients, likely proportion of patients who are capable of or willing to participate, interview procedures for acutely psychotic or other disturbed patients, access to medical records, if needed, and cooperation of ward staff with research personnel, a Review Committee is likely to question if you are sufficiently prepared to face likely difficulties in data collection in this setting, especially, if you don't have a track record in collecting data previously in this setting.

You should provide definitive letters of agreements to participate in the research from credible authorizing official at each data collection site, e.g. hospital, school system, or social service agency. It is also helpful to provide background data, e.g., demographic or diagnostic data, on subjects typically seen at your data collection sites and address any issues of representativeness of your available sample vis-a-vis the general population of interest, e.g. the type of patients typically seen at a psychiatric hospital from which you are collecting data versus characteristics of the general population of psychiatric hospital inpatients.

(2) Respondent burden - The amount of time required to collect data from each research participant should be estimated and carefully considered both with respect to

inconvenience, fatigue, and stress to the respondent and with respect to the quality of data obtainable from overburdened respondents. In general, data collection procedures that require up to 2 hours of respondents' time, especially using self-report measures, are not likely to raise concerns about respondent burden, unless some of the measures seem superfluous or unmotivated. Data collection procedures that require 6 to 8 hours or more of a respondents' time are almost certain to be examined for justification. Such lengthy data collection procedures can create the following concerns: (a) it is often symptomatic of design and analysis problems, esp., inclusion of too many variables in the research, a non-specific data exploration ("fishing expedition") approach, invalid data from measures due to respondent fatigue, and over-reliance on one data source, e.g. self-report data; and (b) such lengthy procedures often require more than one subject session, which can increase problems of subject attrition, or problems with missing data if a subject does not complete all the measures.

Effort should be made to reduce excessive respondent burden by thoughtfully selecting the minimum number of instruments to achieve research goals, use of short forms of research instrument, if they provide adequate information for the research purposes, and use of more than one information source. It is usually worthwhile to focus more data measurement time on the important constructs in the research to ensure their adequate measurement and less time assessing less important constructs. You should avoid the mental set of collecting data because some measures are "usually" collected in research in your area, if such data is not useful in your conceptual framework or will not help you interpret your data, or because the data "might be interesting." Some widely used clinical assessment instruments (e.g., the SCID, DIS, and MMPI) are among the most time-burdensome and aversive to subjects. You should carefully consider whether you really need information on all diagnoses or all clinical subscales from subjects or whether it would be better to use modules or a subset of clinical scales or shorter, more focal assessment instruments in the research. In addition any physically or emotionally aversive procedures, e.g., drawing blood for serum assays or questioning about recent victimization experiences, should be carefully justified from both the human subjects protection perspective and the perspective of the willingness of subjects to participate in the research and value of the data thus collected.

(3) Reactivity of measures. This can refer to either change or distortion of a subject's response due to his/her awareness of being assessed (e.g., responding in socially desirable ways) or to interaction among multiple instruments being collected from the same subjects (e.g., subject introducing more consistency into his reports of different behaviors than would exist if the behavior were assessed directly). The order of administration of research instruments can be important if it introduces error in the form of sequential or carryover effects in subjects responses. This can be particularly a problem with lengthy data gathering sessions or if types of measures are grouped according to content during the data gathering procedures.

(4) Recruitment and training of data collectors - Consideration should be given to the level of experience and sophistication required of data collectors and the amount of training required to ensure the adequacy of data collection. Data collection procedures

requiring complex clinical judgements, involving development of rapport with subjects, or requiring interviews about sensitive topics are likely to require recruitment of data collectors with more training, experience, or personality qualities than is commonly available. Qualifications of these types of data collectors and recruitment procedures to obtain such individuals should be specified. A training plan to ensure adequate data collection procedures by data collectors and to assess the adequacy of the training procedures should be described. Some involved clinical assessment instruments or procedures (e.g., the SCID, the DISC, or the Adult Attachment Interview) require specialized off-site training given by the assessment procedure developers. Such specialized training should be programmed into the proposal's budget and time line.

(5) Supervision of data collection - This may become an issue when data is being collected at multiple sites, when the Principal Investigator is not at the site where data is being collected, or for complex data collection procedures (such as complex coding, administration of complex assessment, or administration of complex interventions) in which variation between data collectors or variation over time in the same data collector may be significant. Detailed description of procedures to adequately supervise and monitor data collectors in these situations should be provided.

(6) Blindness of data collectors - To avoid experimenter bias effects, data collectors, to the extent possible, should be blind to characteristics of subjects that play a role in the research hypotheses of the study, such as group membership, e.g. abused versus control status, or assignment to treatment group, e.g. treatment versus placebo group. If data collectors cannot be made completely blind to subject characteristics, e.g. a questionnaire on contacts with Protective Services is included in the data gathering instruments, it is often possible to plan data collection procedures to minimize the influence of potential data gathering bias, e.g. have the interview on Protective Services contacts be the last data collected from the subjects during the interview session.

COMMON WEAKNESSES IN THE DATA COLLECTION PROCEDURES IN PROPOSALS

1. Failure To Clearly Specify How Data Will Be Collected: Not clear which measures different participants will provide in the study, when they will be assessed; which measures will be collected at different time points in longitudinal studies; how related participants (e.g., spouses, teachers, friends) will be recruited or tested in multiple informant studies, especially when some participants might not have available co-informants (e.g., "spouse" for unmarried participants, "teachers" for school dropouts)
2. Reliance On A Single Data Source: Collection of all data from a single informant or rater; single measure for complex or important constructs

3. Respondent Burden: Administration of lengthy set of measures to a participant in either a single session or multiple sessions, especially, to children or others with limited tolerance for lengthy data collection sessions
4. Reactivity Of Measures: Collection of multiple measures or ratings from a single informant which might interact; respondents remembering previous responses in repeated measure studies; respondents sharing information on responses during course of a study
5. Failure To Specify Training Procedures For Data Collectors: Failure to specify procedures to ensure agreement between multiple raters or coders; no procedures specified to prevent rater or coder drift during course of data collection or coding; no training procedures specified for data collection procedures requiring special training or expertise (e.g., standardized interview procedures, such as the DIS, DISC or DICA or "attachment" assessment using standardized paradigms)
6. Failure To Specify Procedures To Prevent Experimenter Bias From Affecting Research: Failure to specify procedures to ensure blindness in raters or data collectors and lack of bias in data collection procedures, in data reduction, and in data analysis; use of procedures that convey biasing information (e.g., group membership or social stigma) to informants (e.g., asking teachers or clinicians to report on emotional problems of children identified to them as having been abused)

Statistical Analysis

There are four types of statistical data analytic tasks that are commonly conducted in behavioral research: (1) group comparisons using such techniques as multiple comparison procedures, univariate and multivariate analysis of variance, and logistic regression, (2) describing relationships among variables using such techniques as linear regression, multiple regression, generalized linear models, principal component, factor, and canonical correlation analysis, and path, linear dependency, and structural equation models, (3) accounting for variation within samples using such techniques as random effects and hierarchical regression, subgroup analysis, and cluster analysis, and (4) describing change over time using such techniques as repeated measures ANOVA and MANOVA, growth curve analysis, survival analysis, and stochastic models. In addition, statistical techniques are used to assess the reliability and validity of measures and to assess characteristics of the measurement space (e.g., the dimensionality of a measurement domain). In more complex research studies these data analytic tasks are usually combined, for example, a study of the differences in the interrelationships among a set of variables in different groups over time.

Specific statistical techniques are used to describe data or test hypotheses or conceptual models in the research. The statistical analysis must be spelled out, not just outlined. In

particular the statistical method used to verify each specific hypothesis or expected relationship among variables should be described. You should be specific about which particular type of analysis you intend to use for each research question or hypotheses. Reviewers frequently object to data analysis plans that list general techniques to be used in the data analysis (e.g., "logistic regression" or "factor analysis") without specifying which hypotheses and/or particular variables will be analyzed by such techniques, and they also may object when an applicant indicates more than one alternative data analytic strategy or procedure without indicating which one he/she

actually intends to use or how a determination will be made about which one to use. If you have multiple measures of the same construct (e.g. child-reported, parent-reported and teacher-reported measures of child aggression) or multiple measures in the same conceptual domain (e.g various dimensions of family functioning), you should specify how you plan to use, reduce, or combine these multiple measures.

Certain statistical procedures are generally accepted as appropriate for analyzing certain types of data, so that their use is not questioned. If you propose a type of analysis which is not frequently used in your research field or is an extension or deviation from accepted practice, you have to justify your departure from standard practice. For example, some procedures such as ANOVA, MANOVA, multiple regression analysis, factor analysis, and discriminant function analysis, are considered standard procedures with certain types of data and are not usually questioned as long as you specify what variables you are analyzing. On the other hand, other procedures such as cluster analysis, which is considered a group of related procedures rather than one standard procedure, or structural equation modelling, which requires specification of a path model or competing models and has different types of estimation procedures depending on the type of data, need more justification and specification.

Lately, review committees have been more insistent on requiring **subgroup analyses** in their review of data analyses plans. Often such subgroups are readily identifiable (e.g., gender, age, and ethnic groups), but other times less obviously so (e.g., different subtypes of patients with the same psychiatric diagnosis). Again, sometimes subgroup analyses are fairly straightforward (e.g, analysis of variance with subgroups as factors or multiple group comparison approaches), but other times not as straightforward (e.g., separate regressions within different subgroups, hierarchical regressions with "significant" categorical variables, or "random effects" longitudinal models). You should be aware, then, that reviewer are increasingly viewing global analyses of heterogeneous groups as relatively unsophisticated and you should anticipate this potential criticism by paying some attention to subgroup analysis in your data analytic plan.

Unless the statistical analysis is very simple and straightforward, it is advisable to obtain a statistical consultant to help you formulate the statistical analysis for the proposal and to be included as personnel or as a consultant on the grant. If you don't understand the various assumptions involved in formulating and using a particular mathematical/statistical model underlying your proposed statistical or measurement procedures, you need a statistical consultant. On the other hand, we sometimes see

proposals in which it appears that a statistical consultant wrote the data analysis section and someone else wrote the rest of the application - in that, the data analysis section doesn't mesh with the rest of the application or is far more sophisticated than the rest of the proposal. Rather than surrendering control of the data analysis part of the research to a statistical consultant, you should work closely with that person and understand at least conceptually what type of analysis the consultant is proposing so that you can judge its appropriateness for your data.

I think it is a good procedure to actually write out the mathematical form or model that underlies your data analysis, for example, in the form of a (multivariate) analysis of variance linear model, a (multivariate) linear regression model, or a structural equation or path model. You should examine this model to see if you have adequately described (a) all the dependent and independent constructs in the model, especially the reliability and validity of the measures you have chosen to represent these constructs, (b) the relationships among these constructs as specified in your research hypotheses and as justified in your review of previous research and formulation of a conceptual basis for the research, and (c) the statistical power of the data analysis in terms of the number of subjects needed in the various research groups and the ratio of number of variables to number of subjects in various analyses. If you use a repeated measures design, e.g. in a longitudinal or an intervention study, you should specify both a conceptual model of expected change and a plan to analyze change over time or occasions.

You should know that most review committees have members who are quite knowledgeable about measurement, sampling, and statistical methodology. Because of the relatively nonmathematical training of behavioral scientists and the wide availability of packaged computer programming of sophisticated statistical procedures, we often see proposals indicating, with little detail or justification, use of mathematically complicated statistical procedures that the proposer probably does not clearly understand, especially, in regard to limitation or restrictions on the applicability of these techniques. As one example, we have been seeing lately proposals to use LISREL modelling to "discover" relationships between a large number of variables. The estimation procedures used in LISREL have a number of restrictions on distributions of variables and number of subjects and variables in the modelling processes; moreover, it is intended to be a confirmatory procedure rather than a discovery procedure, requiring that you specify your model ahead of time rather than expect LISREL to "discover" it. When the experts on the committee see such a proposal, they rightly feel that proposal's data analysis plan is inadequate. Therefore, if you do propose to use such relatively sophisticated statistical analyses in your proposal, you should have an expert statistical consultant help you formulate the data analysis procedures or, at least, review the procedures you are proposing for their applicability to your data set.

Some well-written, relatively nonmathematical statistical references that may be helpful in understanding the conceptualizations underlying statistical techniques are the following: for data description John Fox and Scott Long (Eds.) Modern Methods of Data Analysis (Sage, 1990), for regression analyses John Fox Applied Regression Analysis, Linear Models, and Related Methods (Sage, 1997), for analysis of variance and

multivariate analysis of variance Richard J. Harris ANAOVA: An Analysis of Variance Primer (Peacock Publishers, 1994) and D. J. Hand and C. C. Taylor Multivariate Analysis of Variance and Repeated Measures (Chapman and Hall, 1987), for design and analysis of comparative treatment studies Joseph Fleiss The Design and Analysis of Clinical Experiments (Wiley, 1986), for design and analysis comparing extant groups, e.g. victims versus controls, Paul R. Rosenbaum Observational Studies (Springer-Verlag, 1995), for categorical data A. Agresti Categorical Data Analysis (Wiley, 1990), for multivariate methods James Stevens Applied Multivariate Statistics for the Social Sciences. Second Edition. (Erlbaum, 1992), Lawrence G. Grimm and Paul R. Yarnold (Eds.) Reading and Understanding Multivariate Statistics, and Ram Gnanadesikan Methods for Statistical Data Analysis of Multivariate Observations. Second Edition (Wiley, 1997), for path and structural equation models John C. Loehlin Latent Variable Models. Second Edition. (Erlbaum, 1992), and for analysis of longitudinal data Linda M. Collins and John L. Horn (Eds.) Best Methods for the Analysis of Change (American Psychological Association, 1991) and James Dwyer, et. al. (Eds.) Statistical Models for Longitudinal Studies of Health (Oxford U. Press, 1992).

Although often a combination of quantitative and qualitative analysis in a research study is viewed as a strength in a proposal, a study proposing only qualitative data analysis is not as readily accepted by review committees as are quantitative methods of analysis and require further justification and description of how data will be collected and analyzed. What the review committee will be concerned about is that qualitative methods are too "subjective", so you must convince the committee that the type of analysis you plan to perform is "objective" enough so that the obtained results could be replicated by another investigator. Often applicants who propose to use qualitative techniques do not describe these techniques in any detail ("the data will be 'coded' " or "thematic analysis of the interviews will be conducted") so that their appropriateness and likely replicability can be judged. If you propose to use qualitative techniques, I would advise that you consult and cite the large literature on qualitative techniques for procedures that have been standardized or otherwise validated. In particular Sage Publishers has published a veritable library of recent books on qualitative methodologies (see, for example, Crabtree, B. and Miller, W. Doing Qualitative Research, Sage, 1992; Denzin, N.K. and Lincoln, Y.S. Handbook of Qualitative Research. Sage, 1994; Creswell, J.W. Qualitative Inquiry and Research Design, Sage, 1998).

Review committees usually require that you perform some type of statistical power analysis for your research design to insure that the sample sizes you propose are sufficient to adequately detect group differences or relationships among constructs (refer, for example, to Jacob Cohen Statistical Power Analysis for the Behavioral Sciences, Second Edition, Academic Press, 1988, Helena Kraemer and Sue Thiemann How Many Subjects?, Sage, 1987, or Ralph O'Brien and Keith Muller Unified power analysis for t-Tests through multivariate hypotheses In Lynn K. Edwards (Ed.) Applied Analysis of Variance in Behavioral Science, Marcel Dekker, 1993). There is a very large and growing literature on methods to calculate statistical power for particular types of statistical techniques and the factors that affect statistical power. Statistical power can often be increased in a study by changes in measurement (e.g., a more precise or accurate measure

of a construct), in design (e.g., matching or stratification) or in data analysis (e.g., use of covariate analysis or a restricted set of multiple comparisons). Statistical packages often include programs or macros to conduct power analysis for particular statistical techniques (e.g., the %POWER and UnifyPow macros for SAS and the SamplePower program for SPSS). A number of power and sample size programs and calculators for types of statistical procedures are available as freeware through the Internet. Much of this is available at a website on power analysis, <http://www.mpl-pwrc.usgs.gov/powcase/powlinks.html>. For complex designs expert statistical consultation should be obtained on calculating the statistical power of a statistical test.

For a number of reasons, often a "power analysis" is not feasible for the actual data analysis you are going to perform. For example, to compute a "power analysis" for a multivariate analysis of variance design requires *a priori* knowledge not only of the variance-covariance matrix of the groups by dependent variables design matrix, but also the "alignment" of the multivariate group means in the multivariate measurement space. Faced with the difficulty of computing a "power analysis" for the actual (complex) statistical analysis to be used in the research, applicants typically adopt one of 3 strategies: (1) compute a power analysis for a similar, though simpler, analysis than the one they are actually interested in (e.g., compute a power analysis for a univariate analysis of variance as opposed to a multivariate analysis of variance); (2) compute a power analysis for the proposed statistical analysis under a set of simplifying assumptions that allow the power to be calculated (e.g., James Stevens in his Applied Multivariate Statistics for the Social Sciences, Second Edition has published tables of statistical power for multivariate analysis of variance using a set of reasonable simplifying assumptions for behavior research); or (3) refer to references in the statistical literature which present power tables for particular complex statistical procedures; these are often based on computer simulation studies of the statistical procedure on different types of hypothetical data sets (e.g., Vonesh, E.F. and Schork, M.A. (1986) Sample sizes in the multivariate analysis of repeated measures. Biometrics, 42, 601-610). In addition to these guides to sample size selection, I, personally, advocate *a posteriori* computation of power for making inferences with particular statistical analyses using data on sample variability obtained from the collected data, although some statisticians object to this procedure because it is too sample dependent.

Regardless of the type of "power analysis" you offer to the review committee to justify the adequacy of your proposed sample size, there are a number of "rules of thumb"-type considerations that review committees use in evaluating the adequacy of the proposed sample size: (1) One consideration concerns the number of subjects in the smallest significant cell in the design. Thus, if you propose to compare the effects of 2 different interventions with males and females from 2 different ethnic groups, the review committee is not as likely to focus on the total number of subjects recruited, but on the number of subjects in the smallest cell about which you wish to make inference (e.g., here, the differential effects of the 2 treatments in females in one of the ethnic groups). This issue of cell size can become particularly problematical in comparative or field studies in which the natural distribution of subgroups or ability to recruit subjects

representative of certain cells in the design is vastly different (e.g., in studies of the differential impact of types of abuse in different sex and ethnic groups, recruiting Black males who have been sexually abused). (2) A second consideration concerns the ratio of number of variables collected to the number of subjects. If you collect 30 variables from each of 30 subjects, the review committee is liable to feel that, because of the common source variance and the large number of variables collected, there is a high probability of "discovering" spurious relationships among variables in the research. (3) A third consideration concerns the number of subjects typically recruited in studies in the area. For example, if, in studies of nationally representative samples, 2000 subjects are typically included and you propose a study with only 500 subjects, the review committee is liable to feel your sample is too small. In contrast, if typically, in the proposed research area, comparison of different treatments is based on studies using 20 subjects in a group and you are proposing a study with 50 in a group, your sample size will look large to the committee. In considering the adequacy of the sample size you propose to use in the study, you should also pay attention to these considerations, because you should remember that the review committee members will rarely, if ever, check your power calculations (no matter how you finesse the computation of the power for your study) and that, no matter what you calculate as the power of the design, if you only have 5 subjects in a cell or twice as many variables as subjects or half the number of subjects as in typically used in research in your area, the review committee will have serious concerns about the adequacy of the sample size and this could be fatal or near fatal to the chances of your proposal receiving a highly favorable evaluation.

COMMON WEAKNESSES IN THE DATA ANALYSIS PLAN OF PROPOSALS

1. Failure To Describe How Specific Hypotheses Will Be Tested Or Analyses Will Be Conducted: Failure to specify the specific variables to be used in a hypothesis test or data analysis and the specific statistical or other data analytic technique to be used; reference to general statistical techniques without specifying which variables will be used (e.g., "MANOVA will be used");
2. Lack Of Congruence Between How Data Will Be Analyzed And The Aims of the Study: Specifying a statistical procedures that does not adequately test the major hypotheses of the study; lumping multiple measures together in a multivariate procedures instead of focusing on specific hypothesis testing; specification of sophisticated statistical procedures which are of limited use with data collected, especially when recommended by a statistical consultant who is not familiar with the substantive research area
3. Use of Complex Statistical Procedures That Are Not Well Understood By The Investigator: Failure to consider the limitations or prerequisites for use of statistical procedures (e.g., distributional assumptions, required structure of covariance matrices, level of measurement requirements)

4. Use Of Simple Statistical Procedures When More Complex Procedures Are Necessary Or Commonly Used: Use of multiple univariate tests when multivariate tests are more appropriate; adding together items or measures to obtain a single score when a psychometric item or scale analysis is appropriate;

5. Use Of Statistical Procedures Which Are Not State Of The Art: Use of antiquated statistical procedures learned 20 years ago in graduate school rather than current statistical methodology, especially the most sophisticated procedures used in the proposed research area

Two very general issues which affect more than one aspect of the research proposal namely, the "logic" of the research and dealing with "multiplicities," will be briefly discussed because they often have important repercussions for the review of applications.

The Logic of the Research

This term denotes the extent to which there is a logical "connectedness," specificity, and interrelation among the various components of the applications. Thus, "good logic" would encompass a theoretical framework or specific hypotheses developed in the literature review and review of preliminary results; the theoretical framework or hypotheses in turn determine which variables are chosen to be measured in the research and determine the type of research design that allows the conceptual model or specific hypotheses to be tested; and the theoretical framework or hypotheses also allow the definition of appropriate study populations to test the research model or hypotheses; and the data analysis plan tests the specific hypotheses or model developed in the proposal using the measures collected in the research; and the research as a whole is relevant to accomplishing the stated aims of the research.

Problems with the "logic" of the research include: (a) positing or adopting a very general conceptual model that is too vague or general to allow the statement of specific hypotheses or a testable empirical representation of the model or to be operationalized by adequate measures, (b) inclusion of important constructs in the theoretical framework for which no justification or ground have been developed in the literature review or review of previous research, (c) inclusion of constructs in the conceptual framework which are not measured by any instrument or procedure in the data collection phase or, contrarily, inclusion of variables in the data collection phase which are not discussed in the development of the conceptual framework (this is particularly prevalent in the inclusion of popular measures that "everyone uses," e.g., the Achenbach Child Behavior Checklist, the DISC, and the SCL-90), and (d) failure in the data analysis section to describe specifically how each hypothesis or the model will be tested.

Some devices that might improve the "logic" of the research are:

(1) Create a table in the "Experimental Design and Methods" section that lists all of the important constructs in the research and the specific measures collected in the research

that operationalize these construct (possibly with data or references to the psychometric adequacy of the particular measures). Examine this table to judge how well each construct was discussed, defined, and justified in the "Background and Significance" and "Preliminary Studies" sections when the theoretical model or specific hypotheses were proposed and to judge the adequacy of the operationalization of the construct by the specific measures chosen. Compare this table to the measures actually collected in the data collection procedures for "stray" variables that should be discussed and incorporated into the theoretical framework or that should be "backgrounded" or deleted from the study.

(2) State the specific hypotheses that are the aims of the research or present a representation of the theoretical model, e.g., in the form of diagrams, charts, or equations, at the conclusion of the "Background" or "Preliminary Studies" section, then restate the hypotheses or redescribe or refer to the representation of the model at the beginning of the "Data Analysis" section of the "Experimental Design and Methods" section and discuss data analysis methods to test each specific hypothesis or the model (and possibly alternative models) as a whole.

Dealing with "Multiplicities"

Often various types of "multiplicities" arise in aspects of the research proposal that are ignored or left unanalyzed and that create problems in the evaluation of the research. Some examples of these types of "multiplicities" and their attendant problems are:

(a) multicomponent or multidimensional constructs - Many behavioral constructs have multiple components, dimensions, or manifestations; which of the components, dimensions, or manifestations are focused on in the conceptual framework, the measurement of the construct, and the interpretation of results?

(b) co-morbid problems or diagnoses in clinical populations - Often subjects included in clinical investigations or intervention studies may have multiple problems or meet diagnostic criteria for more than one psychiatric diagnosis (e.g., substance abuse problems in batterers or major depression in PTSD patients); how are such multiple problems or diagnoses to be dealt with in terms of inclusion/exclusion criteria or in treatment protocols?

(c) multiple data sources - There might be multiple measures of a construct or multiple reporters of behavior; how are these multiple sources to be combined to measure a construct?

(d) multiple data collection points - What is the justification for the multiple data points, the interval between data collection points, and specifically how will this time dependent data be analyzed?

(e) multiple samples - If data is collected from more than one site or referral source, how comparable are the samples and, if they are not, how are their results going to be combined or compared?

(f) multiple groups - If multiple comparison or control groups are used in the research, what is the rationale for choice of comparison and control groups?

(g) multiple methods of data analysis - There is often more than one approach or method of analyzing the same data (e.g., different types of "cluster" analysis, discriminant function analysis vs. logistic regression, or different types of "analysis of variance with unbalanced data," such as sequential vs. unique sum of squares), what is the rationale for choosing one method of data analysis over another?

There are different ways of dealing with these multiplicities, but often applicants do not address these, leaving the reviewer hanging about how these problematic issue may be addressed. If you do not have a clear plan and/or rationale for ways of handling the types of multiplicities in your research, you should do a careful review of the literature in your field, consult with knowledgeable colleagues or expert consultants about the best ways to deal with such multiplicities, and include these plans or rationales in the proposal.

Human Subjects

An important consideration is the adequate protection of human subjects in the research. Research funded by NIMH that involves contact with human subjects must conform to Federal regulations regarding conduct of research with human subjects (primarily Code of Federal Regulations, 45 CFR 46 and amendments). In particular the applicant institution must meet certain requirements for conducting an institutional review of the procedures proposed in the research to protect human subjects from risks associated with participating in the research project. Most large universities, medical school, and research institutions have in place an agreement with the Department of Health and Human Services concerning conducting such reviews by an Institutional Review Board (IRB). IRBs must meet certain requirements concerning their composition and operation. If such an agreement is in place, the institution is issued a Multiple Project Assurance allowing the IRB at the institution to review research projects submitted for DHHS funding. Usually this review takes place prior to submission of the application and is indicated on the face page of the application by listing the Multiple Project Assurance agreement number assigned the institution (an M number) and the date of the IRB review. IRB reviews must be current within a year of both review and funding.

If your institution does not have such a standing review board approved by DHHS, it must negotiate with the Office for Protection from Research Risks (OPRR) at the National Institutes of Health to establish such a board to review your project. When such an approved board is in place and other requirements are met, OPRR will issue a Single Project Assurance for your project (designated by an S number). Negotiating such a Single Project Assurance can be quite lengthy and burdensome, but your application can not be funded without either a Multiple Project or a Single Project Assurance review. If

your project will need a Single Project Assurance, I would advise you to consult with OPRR regarding their requirements as you are preparing your application. Although you can initiate procedures to obtain a Single Project Assurance prior to IRG review, in general it makes more sense to apply for such an assurance only if your funding prospects are good.

Some research is exempt from human subject protection requirements. Be aware, however, that the applicable Federal regulations granting exemptions from human subjects protections and DHHS interpretation of such regulations are very narrow. Exempted research usually falls in the category of research on extant records or research on data routinely collected in public agencies, educational institutions, or clinics. Anonymously collected survey data is also exempted **except where the data collected is on sensitive topics (e.g., on sexual or violent behavior)**. In general, any research in which data is collected directly from human subjects by the research team is **not** exempt from requirements for human subjects protection. If you or your IRB considers your proposed project to be exempt from human subjects protection requirements, I would advise you to consult with either program staff or the Office for Protection from Research Risk concerning whether your claimed exemption is acceptable.

Some researchers assume that obtaining approval for a research project from their institutional IRB is sufficient to meet human subject requirement for research supported by NIMH. Such is not the case. We find often that IRBs vary greatly in the criteria they use to evaluate provisions for human subject protections, sometimes underprotecting subjects and sometimes overprotecting subject from minimal risks. Human subjects protections will also be reviewed by the review committee and by NIMH program staff for their adequacy. Sometimes this can create difficulties for an applicants when there is a conflict between human subjects protection requirement of the IRG or program staff and the institutional IRB; changes in the human subjects protections procedures made by an applicant in response to an IRG review or program staff requirements must be reapproved by the institutional IRB.

There are four main concerns in research with human subjects: (a) obtaining voluntary consent; (b) informing participants of potential risks; (c) attending to potential adverse effects of the research; and (d) safeguarding the confidentiality of information obtained from participants. A research proposal in which there are substantial concerns about safeguarding the well-being or confidentiality of subjects will not be approved by the review committee nor funded unless corrected.

Often applicants do not realize some of the intricacies involved in human subject protection procedures. This is particularly true if you are conducting research with certain populations (e.g., children, adolescents, mental patients, prisoners, victims of violence, disasters, or other traumatic events, delinquents or criminals who are not incarcerated, and child or spouse abusers) or on certain issues involving sensitive, stressful, deviant, or illegal behavior (e.g., child abuse, spouse abuse, rape, community violence, deviant or illegal sexual behavior, and involuntary civil commitment). In particular, if your research involves actual individuals from such subgroups or who are a party to such behavior, a

statement, that there is "no risk" from participation in the research and, thus, no procedures for protection from research risk need be specified, will be flatly rejected by the review committee. In such cases I would recommend talking to knowledgeable colleagues, especially researchers who have or have had Federal funding for research in the field, or NIMH staff about protocols and the best available human protection procedures being used in the research field. It is also often helpful to establish liaisons or specific agreements with agencies that might play a role in providing services to participants or that participants might be referred to, or that might have some jurisdiction over participants (e.g., clinics, child protective service agencies, courts, or parole agencies).

Informed Consent

Obtaining informed consent is a process consisting of the following components: (1) apprising potential participants of the general nature of the research, the type of information to be collected, and potential risks clearly enough so that potential participants can decide whether or not they wish to participate in the research; (2) ensuring that the participants understand the explanation of the research procedures and risk well enough to consent to participate in the research; and (3) obtaining voluntary consent to participate. Although typically informed consent is documented by asking participants to sign a voluntary consent form, it is important to realize that informed consent is not a document, but rather a process whereby the investigator attempts to guarantee that potential participants have enough knowledge of the research to truly make informed decisions about whether they want to participate in the research project. The consent form that will be used in the research and a description of the process used to obtain voluntary consent should be included in the application and will be reviewed by the review committee.

The explanation of the research should be targeted toward the level of understanding of potential participants (e.g., age, education, emotional stability). Brief, abstract, vague, or overly general accounts of the research are not acceptable as providing adequate information to potential participants on which to base their voluntary consent. This is especially true if the only information is provided in a written consent form. Particular aspects of the research which might be problematic for some participants (such as sensitive topics the applicant will be asked about, burdensome or stressful assessment instruments, or invasive or intrusive procedures) should be clearly identified and explained to the subject together with their options for participating or withdrawing from the research or particular procedures. Research involving partial disclosure or deception are particularly problematic (e.g., not fully disclosing or misleading participants about the goals of the research or design of the study, for example, that the subject might be randomly assigned to a control group during a treatment study). Potential participants should be informed of the expected length of data collection sessions and the number of sessions they will be asked to participate in.

Depending on the nature of the research and the study population, there can be some problematic issues concerning obtaining voluntary consent. In research involving

children and adolescents, the researcher is generally required to obtain voluntary consent to participate from the parent or legally responsible caregiver and to also obtain consent or permission from the child or adolescent. Depending on the age of the child and his/her cognitive and emotional maturity, the consent procedure for children and adolescents may be similar to that for adults ("consent") or more in the nature of an explanation of what is involved in the research and the child's options for participation in the research as a whole or in particular procedures ("assent"). In most jurisdictions children and adolescents below a certain age may not solely give permission to participate in a research project; consent also has to be from a legally responsible caretaker. In some instances this can be quite problematical (e.g., runaway or throwaway children or children in foster care). In general "passive" consent procedures (i.e., where an individual or a child's parent is assumed to be willing to participate in the research unless they specifically decline to participate) are not acceptable.

Any appearance of coercion, no matter how subtle must be avoided (e.g., unusually large subject payments, a perception that receipt of services at an agency is dependent on participation in the research, or parent coercing reluctant children to participate in research, especially in a study offering a substantial subject payment to parents). This is particularly true in research involving "captive" populations, such as children, crime victims, mental patients, and prisoners. Participants should be informed of their freedom to withdraw from the research at any time and to refuse to participate in any particular aspect of the research that makes them uncomfortable.

Informing Participants of Research Risks

Any potential risk from participating in the research must be clearly explained to the potential participant before consent is obtained (e.g. if information is obtained from a family member that could lead to an abuse report, such reporting requirements must be explained to family members prior to obtaining consent). In research with children that might involve reporting of child abuse it is important that you be aware of current State statutes mandating child abuse reporting and ,if possible, to develop a working relationship with child protective service agencies. This will inform you of criteria that the agency is likely to use for investigating reports (sometimes investigators might report instances of alleged abuse that an agency regards as so minor that it would not even conduct an investigation) and of services that might be available or are likely to be offered to abusive families (if any, and which might, for ethical reasons, need to be supplemented by either direct services offered by the research team or through referral to service providers outside the protective service system). You should be clear about the exact procedures that will be used when allegations of abuse are uncovered during the course of the research (e.g., who is responsible for verifying that an incident is reportable, who is the incident reported to (e.g., directly to protective services, to a school or agency providing subjects). You should be quite leary of "investigating" potentially abusive episodes yourself since most researchers do not have the legal mandate nor the experience to conduct an adequate investigation of abuse.

There also may be limitation to confidentiality of data obtained from subjects (e.g., data obtained from adolescents may be accessible by their parents unless some waiver is obtained from parents). Such potential limitations to confidentiality should be thought through and addressed either with additional protective procedures or by advising potential participants of the limitations to confidentiality. It may be advisable to seek legal advice about such limitations and additional protective procedures.

Minimizing Risk

Provisions must be made for potential adverse reactions due to participation in the research either of informants or by others directed towards the informant; this is especially true of adverse emotional reactions of participants (e.g. from crime or abuse victims) and potential retaliation by the perpetrators for abuse reporting by assaulted spouses or abused children. These provisions usually include adequate staff training to detect such reactions and procedures to provide clinical services for problems related to the research or identified during the research, for referral to adequate helping resources, or for adequate physical safety. In research on sensitive issues, involving invasive or intrusive procedures, or with high risk populations, debriefing and followup procedures should be planned to assess whether participants experienced any discomfort or difficulties from participating in the research immediately after participation or after some delay. Adequate provision should be made to address such discomfort or difficulties.

Some studies, especially studies involving biological measures, require that subjects be taken off medication for a period of time before the subject is assessed to prevent medication effects from influencing the assessment. The implications of withdrawing medication should be carefully weighed in terms of the health and well-being of subjects and the course of the condition for which medication was prescribed. The condition of such subjects should be carefully monitored.

At a minimum safeguarding of confidentiality requires numerically coded data forms and locked files. A DHHS Certificate of Confidentiality can be applied for through NIMH when data collected might be subject to subpoena. Depending on the type of research or the study population, limitations on confidentiality must be clearly explained to the subject (e.g., mandated child abuse reporting, parents' access to children's or adolescents research data, *Tarasoff* limitations). In particular, NIMH does not regard the Certificate of Confidentiality as superceding mandated State reporting requirement for child abuse. In the event there is a conflict about release of confidential research data, confidentiality provisions spelled out in the consent form are often as important as the Certificate of Confidentiality. In situations in which there is a possibility of legal subpoena of research data (e.g., in research with recent rape victims), I would suggest consultation with your organizational legal counsel about the wording of consent forms and research procedures to follow that would better guarantee the confidentiality of the research data collected.

Human Subjects Issues

NIMH has established a special human subjects review panel to scrutinize problematic human subjects issues in applications and funded research. Such problematic issues may be referred to this panel by IRGs, NIMH staff or the Advisory Council. The panel may assess the adequacy of human subjects protection in addition to reviews by IRBs and IRGs and require clarification of human subjects protection protocols or additional protections in such protocols. There are a number of human subjects issues that arise relatively frequently across mental health research studies that require attention in the human subjects protocol of the study. Among these are:

Child abuse reporting. All states have legal requirements that certain professionals report known or suspected incidents of child abuse. States vary in the types of professionals and/or nonprofessionals required to report abuse, the standards of evidence for reports, the degree of recency of reportable abusive behavior and provisions for immunity from lawsuits for good faith reporting. In addition to legal issues, there are also ethical issues of dangerousness to children and need for treatment for past abuse and potential for harm to family integrity that play a role in abuse reporting. In general, NIMH does not allow use of a Certificate of Confidentiality to avoid reporting of child abuse. The likelihood of abuse reporting depends on the type of study and the population being recruited for the study. Thus, studies that inquire about experiences of traumatic events in children from distressed families or neighborhoods is highly likely to uncover incidents of unreported child abuse. Studies of normal developmental processes in children from non-distressed families are unlikely to uncover cases of child abuse. In studies that have a significant probability of reporting child abuse, such reporting can result in significant negative consequences for the child or family and, thus, must be highlighted in the consent procedures as a potential significant risk of participating in the research. Because of the complex legal and ethical issues involved in child abuse reporting, it is suggested that investigators who conduct studies that might result in child abuse reports consult with NIMH staff or recognized child abuse researchers about protocols for handling child abuse reporting. Other types of studies that involve issues of dangerousness to self or others also require attention to safety, referral, or reporting of research participants (e.g., studies that uncover spousal assault or abuse).

Need for Treatment. The identification of research participants who are in need of mental health services (e.g., suicidal individuals or individuals with significant mental health problems) usually require that such individuals be provided or referred to appropriate treatment. The level of referral offered depends on the nature of the study and the degree of risk of the study population. Thus, a study of schizophrenics in an inpatient setting should have in place a system of monitoring and collaboration with treatment providers in the inpatient setting, whereas in a study of daily hassles in a community sample it might be sufficient to be able to inform individuals with mental health problems of available community treatment resources. If referral to existing treatment programs is offered, it is far better to have agreements with such referral programs that they will accept or respond to the referrals that you make than just to provide a list of referral programs.

Withholding Treatment or Medication. Studies that require withholding or withdrawing individuals from treatment are quite problematic from an ethical perspective. This is particularly the case when effective treatment that reduce symptoms or alleviate symptoms are known. There exist some rationales for withholding or withdrawing treatment (e.g., when the withholding is quite time-limited and it is known that adverse effects are highly unlikely or when existing treatments are highly ineffective and the knowledge achieved from a study is likely to be highly clinically beneficial) but such action in the context of research is likely to be heavily scrutinized for potential adverse effects on research participants.

Challenge studies. Studies that include procedures that induce symptoms or negative reactions in research participant must provide an exceedingly cogent rationale for the necessity for such research procedures. This rationale must be particularly strong when the procedures are applied to previously nonsymptomatic individuals or to particularly vulnerable subjects (e.g., mental patients or children). Such procedures may include biological procedures (e.g., CO₂ challenge to induce panic symptoms) or psychological procedures (e.g., exposure of combat veterans with PTSD to combat films). At a minimum, human subject protection protocols in challenge studies should include minimization of adverse effects, clinical backup to deal with acute reactions, stabilization and debriefing following the experimental procedures, and follow-up of research participants to monitor potential adverse effects of participation in the research.

Deception. Because of requirements that research participants be apprised of the nature of research goals and procedures and potential risks of participating in a research study in order to give informed consent for participation in the research, it is seldom justified to mislead or withhold information from potential research participants concerning the aims and procedures of the research. Studies which actively use deception (e.g., use of confederates to induce negative emotions in research subjects) are particularly problematic from both a research (i.e., it is often uncertain the extent to which the subject was unaware of the manipulation) and from an ethical perspective (especially, informed consent). Some studies may justifiably present the goals of the research in neutral terms to avoid stigmatization (e.g., a study comparing children at high risk for delinquency with low-risk children may be described as a study of social development in children) as long as the data collection procedures of the study are clearly specified to potential participants. Potential participants in a treatment study who might be randomly assigned to a placebo or nonspecialized treatment should be informed of this aspect of the study design in order for them to make an informed decision about participating as a potential control subject; this is especially important for treatment-seeking individuals.

Sensitive issues for consumers or communities. Research that involves potential negative consequences, adverse outcomes, or stigmatization for research participants, consumer groups, or whole communities must involve the greatest degree of sensitivity by researchers to such consequences, real or perceived. Such sensitive issues include, for example, studies of negative family interactions in families with a member with a major mental health problem (e.g., that might support an inference that the family environment "caused" the mental health problem) or studies of biological factors in problematic

behavior, especially in minority communities (e.g., that might support an inference that minority individuals are genetically determined to exhibit the problematic behavior). Such studies may be proposed by researchers with very little contact with such groups or communities and a great deal of naivete about the social context in which research results may be interpreted. Individuals who intend to conduct research of a sensitive nature are advised to spend considerable time discussing the research plan, potential results and potential adverse interpretations with informed members of the affected groups or communities. A research advisory panel with substantial representation from the groups or communities that may be adversely impacted by results of the research or misinterpretations of the research results is strongly advised.

COMMON WEAKNESSES IN HUMAN SUBJECT PROTECTION IN PROPOSALS

1. Failure To Clearly Describe Consent Procedures: No inclusion of consent form in application when research involves potential risks; failure to specify in the consent form potential risks of participation in the research and of the subject's freedom to refuse to participate or to withdraw from the research
2. Failure To Attend To Potential Risks Of Research Participation: No training plan, selection procedures for interviewers, or clinical backup to deal with potential emotional upset in research on emotionally charged issues; no warning of child abuse reporting requirements for unreported parents participating in the research; no attention to possibility of legal subpoena for subject's research data
3. Failure To Attend To Ethical And Legal Requirements That Apply To Professionals: No attention to ethical and legal requirement for child abuse reporting, duty to warn potential victims of violence, and for need for treatment
4. Failure To Specify Procedures For Confidentiality Of Data: No provision for restricting access to data; no discussion of disposition of videotaped data; violation of confidentiality in obtaining access to subjects or to locating subjects in longitudinal studies
5. Assumption That Institutional IRB Approval Satisfies Human Subject Protection Requirements: Institutional IRB approval does not guarantee approval of inappropriate human subject procedures by a Review Committee nor by program staff

Intervention Research

Research on preventive or treatment interventions has additional research design criteria that are evaluated in a grant review. Intervention research is concerned with establishing the validity of an inference that an intervention effected a change in individuals

(sometimes groups) that resulted in a significant change in some desired outcome. As with other types of research, proposed intervention research should be based on a strong conceptual model. The elements of a conceptual framework for intervention research include: (1) a conceptualization of **what to change**, (2) a conceptualization of **how the intervention changes the target of the intervention**, and (3) a conceptualization of **what factors affect variation in intervention outcome**, and (4) a conceptualization of **how to maintain intervention gains over time**.

The methodology of intervention research is concerned with how to ensure the validity of the inference that the change in outcome resulted from the intervention per se and to eliminate or control alternative explanations for change in outcome (internal validity of the intervention inference) and to establish the robustness or variability of the intervention effect across clients, treatment administration, and settings.

Many of the issues raised in this discussion of intervention research are discussed in greater depth in Alan E. Kazdin, Research Design in Clinical Psychology. Second Edition. (Macmillan, 1992).

Some issues in developing intervention research follow:

Assessment of intervention effects

Assessment of intervention effects should focus on precisely what the intervention is designed to change. Often interventions for mental health problems or dysfunctional behavior are not designed to directly affect these problems or behavior, but to change cognitive, emotional, or behavioral processes that are hypothesized to mediate the problems or behavior (e.g., negative cognitions in depressed individuals) or to modify other types of risk factors associated with the problems or dysfunction (e.g., poverty or drug usage in parents neglectful of their children). Evaluation of intervention effectiveness should assess the mediating or risk factors that are hypothesized to be changed by the intervention. Assessment of the mental health or behavioral outcomes that are the ultimate goal of the intervention is usually also assessed in intervention research. If the intervention positively affects both the mediating or risk factors targeted by the intervention and the ultimate mental health problems or dysfunction, the intervention trial not only demonstrates the effectiveness of the treatment, but also tends to validate the mediation or risk factor model linking the processes modified by the intervention and the ultimate problem or dysfunction. It may be the case that the intervention effectively changes the hypothesized mediating targets, but shows nonsignificant improvement in the ultimate problem or dysfunction for a number of reasons (e.g., difficulty in assessing change in the problem, for example, low-base rate behaviors, the hypothesized mediational model is inadequate, or the change in the mediators is not sufficient to effect major change in the ultimate problem). This might be particularly true when enduring, trait-like characteristics (e.g., self-esteem or personality disorders) are assessed prior to and following short-term interventions.

The investigator should also carefully assess whether outcome measures chosen are likely to be sensitive to change in intervention targets. For examples, some assessment measures may not be discriminating enough to adequately assess change (e.g., instruments with only a few categories, such as instruments that assess whether or not a person meet a diagnosis or not). With complex, long-term or sequential types of interventions, assessment of effectiveness at various critical transition points in treatment is recommended.

Intervention specification

To assess whether a particular conceptually based intervention effected change in targeted outcomes requires that the intervention be specified clearly enough and adequately implemented such that it can be determined whether the research trial adequately tested the intervention approach in the sense that the clients actually received the proposed intervention. This determination is usually based on (1) a clear specification of how the intervention is to be implemented in the form of a manual and or other training material, (2) procedures to adequately train intervention providers in the administration of the intervention, and (3) procedures to assess whether intervenors are actually administering the intervention as designed (they are *adhering* to the treatment plan), possibly with corrective feedback procedures to correct incorrect treatment administration. Related to this is the assessment of *differential intervention fidelity*, which consists of procedures to assess whether therapists are faithfully administering an intervention as intended in the treatment condition, but also are not using the target treatment interventions in other intervention conditions or in the control condition. This might particularly be a problem when the same therapists are administering more than one intervention condition or when therapists in the same setting might communicate or otherwise pick up the target intervention procedures and apply them to the control or other treatment groups. Such contaminations adversely affect the estimation of the efficacy of the target intervention process compared to other interventions or controls.

Additional insight into the effectiveness of interventions can be gained by assessment of the process of intervention. Such assessment can take the form of linking various characteristics of the intervention process to measures of intervention outcome (e.g., characteristics of the intervenor's or client's behavior during sessions). One important form of this type of assessment is an assessment of **dose-response** relationship. This concept, borrowed from pharmacological studies, can involve various measures of the quantity or intensity of intervention received. For example, assessment of dose-response relationship could include the relationship of outcomes to such measures as number of intervention sessions administered, number of intervention sessions attended, completing versus dropping out of treatment, and level of client involvement in treatment (e.g., "homework" assignments completed, active participation in intervention sessions).

Factors Affecting Variation in Intervention Outcome

Such factors can include client characteristics, intervenor characteristics, interactions between client and intervenor characteristics, and intra-session and extra-session

processes and conditions that affect the administration of the intervention or response to intervention. The impact of some of these factors might be either controlled experimentally or assessed during an intervention trial.

Among client characteristics that might affect intervention outcome are severity of presenting problem, co-occurring problems (e.g., alcohol abuse in spousal assaulters), and motivation for treatment (e.g., help seekers versus mandated clients). The level of severity of the presenting or target problem should, particularly, be considered in terms of representativeness of intervention trial subject to the target clinical population, the difficulty of demonstrating clinically significant change in client samples with only mild or moderate levels of dysfunction or impairment, the refractoriness of clients with severe levels of dysfunction or multiple types of difficulties, and the possibility of overestimating treatment effects by using extreme groups. If intervention trial participants are relatively unselected for level of severity of target problems, it is often useful to *post hoc* stratify subjects by level of problem severity to gauge the relationship between intervention effectiveness and problem severity.

Among intervenor characteristics that can affect intervention outcome are intervenor competence, which is often assessed as years of experience or amount of training, intervenor adherence to the intervention protocol, or intervenor bias toward an intervention. Similarly, some clinical research has focused on the match between client characteristics and therapist characteristics, especially on such characteristics as race, ethnicity, sex, age, and personality style.

Many intra-session factors that influence intervention outcome are assessed in studies of intervention process. Among these are development of a "working alliance" between therapist and client, the "quality" and timing of interventions by the intervenor in sessions, and the response to interventions by the client in sessions. Among extra-session conditions or factors that can influence intervention outcome is support for intervention participation by significant others, experience of extra-therapeutic crises and life stressors, and other interventions received outside the target intervention.

Often conditions for administration of treatment are not ideal. Some writers on intervention research distinguish between intervention trials of the *efficacy* of intervention versus trials of the *effectiveness* of interventions. Supposedly, in an *efficacy* trial the researcher can control the selection and assessment of subjects, of therapist, and of the administration of treatment such that the trial is an adequate test of effectiveness of the intervention model under near ideal conditions (in other words, of the internal validity of the inference that the intervention model can effect change in the target of the intervention). *Effectiveness* studies, which often follow *efficacy* trials, then attempt to assess whether interventions that have been shown to be efficacious in controlled trials maintain their efficacy in real-world clinical settings. My own view is that the categorical distinction between *efficacy* and *effectiveness* trials is artificial. Rather I think there is a continuum of control that researchers have over aspects of the intervention (i.e., the

researcher may or may not be able to adequately control various factors in the intervention such as selection of subjects and of intervenors, of assignment of subjects to treatment groups, and of administration of interventions) and that any intervention trial falls somewhere on this (these) continuum (continua). Moreover, I think that the assessment of factors that influence intervention outcome is part of a complete intervention model.

Research designs in intervention research

There are a number of research designs used in intervention research. They can be split into multiple group and single group designs. Multiple group designs include (1) intervention versus control or comparison groups and (2) comparisons of multiple alternative interventions often also including control or comparison groups. Special cases of the latter include comparing groups receiving a standard intervention with those receiving the standard intervention plus additional components, which are usually attempting to determine if treatment efficacy can be enhanced by adding additional intervention components to a known effective intervention, and disaggregation of multifaceted interventions, which usually attempt to determine what elements of complex interventions are actually instrumental in effecting change.

Single group designs are usually regarded as not as strong intervention research designs as multiple group designs because outcomes can not be compared to no treatment or to alternative interventions. Nonetheless, there are occasions in which such designs are necessary and may even be superior to standard multiple group studies. Among single group designs are single subject designs in which treatment efficacy is inferred from synchrony between administration of treatment and change in outcome target from a baseline. Often single case results are aggregated in an intervention study. Another form is single sample pre-, post- designs which measure change in the time period covered by an intervention. Such designs are usually most persuasive when the target of the intervention is strongly resistant to change such that it is unlikely that change in the targeted outcome could result from other factors beside the intervention. Another form of single sample design is when a series of stages of improvement in difficult outcomes (e.g., personality change, recovery from serious trauma) are defined and the subject advances through these stages in synchrony with interventions targeted at outcomes characteristic of that particular stage. Outcomes characteristic of other stages may also be assessed to indicate that improvement is not general, but is restricted to the outcomes targeted for intervention.

Randomization plays an important role in supporting the validity of the inference that the intervention effected change in some outcome measure against alternative explanations. Thus, random assignment of potential subjects to treatment or control groups guards against the possibility that differential effects of interventions or of interventions over control is due to pre-existing differences between the groups prior to intervention or in response to intervention. Similarly randomization can be used as a control over other factors that might effect outcome beside administration of treatment (e.g., random assignment of therapists to different treatments to guard against therapist competence as

effecting outcome). However, it is important to realize that randomization is a design *strategy* to control for nontreatment factors; there are situations in which randomization does not serve this function well (e.g., in which randomly assigned groups become biased by differential attrition during treatment administration) and in which randomization adversely affects valid inference about treatment effectiveness (e.g., when treatment efficacy is maximized by matching treatment with client characteristics).

Ethical issues in intervention research

In addition to the general issues concerning informed consent, confidentiality and human subject protection previously discussed there are additional human subject issues specific to intervention research.

An important issue in informed consent is that if potential subjects are to be randomly assigned to different intervention groups or to interventions and no treatment or placebo groups, they must be informed about this random assignment. This disclosure can significantly impact the willingness of potential subjects to participate in the study particularly if they feel that they might not receive intervention. Other consent issues arise with potential subjects who may be coerced into interventions (e.g, children by their parents or individuals court-mandated for intervention).

Additional human subject safeguards are required in intervention research associated with the issue of the demonstrated efficacy of treatments. So, for example, it might be difficult to justify administering experimental interventions to subjects in place of interventions known to be effective in a study in which the intent is to demonstrate that the experimental approach is even more effective. Similarly, human subject protection usually requires that harmful intervention be terminated as soon as the harm is known and that if a particular intervention is conclusively demonstrated to be superior to other interventions during the course of a trial that the less effective interventions be terminated and the effective intervention be substituted for all subjects.

Resubmissions

The odds are high that your first submission of a grant proposal will not be approved with a fundable priority score on first review, especially if you have not had previously funded research.

To increase your chances of funding you will probably want to revise your application and resubmit it for another review. Currently, resubmissions of previously submitted research applications are allotted an extra month from new application submission dates (i.e., resubmission deadlines for research applications are **March 1, July 1, and November 1**). You are allotted an additional 3 pages at the beginning of the narrative portion of a resubmitted application to explain the changes you have made in the application in revising it. You **must** complete these pages in a resubmitted application. Applications which are not revised (as indicated in these preliminary pages) will not be accepted by the Center for Scientific Review at NIH for re-review and will be returned to

you. NIH has adopted a policy of limiting the number of resubmissions to 2 (i.e., you can submit an original application and then at most 2 resubmissions of the original application) and the time period in which you can revise and resubmit an original application to 2 years (i.e., you can only resubmit revised applications within a 2 year period from the date of the original submission of the application).

Your chances of receiving a fundable approval when you resubmit a revised application are vastly improved provided you satisfactorily respond to all the criticisms raised in the initial review. Usually, when an application is re-reviewed by a committee the same reviewers who initially reviewed the application will be assigned to review the resubmission, provided they are still serving on the committee. In addition to the revised application, these reviewers will also be sent the Summary Statement from the original application by the Executive Secretary of the review committee, and, although they are not sent the original application, they are likely to remember its good and bad features. In general reviewers judge the adequacy of a resubmission according to the criteria of **how well the applicant responded to the issues raised in the reviewers' critiques**.

The best strategy to adopt in responding to criticisms is to determine which issues were identified by reviewers as the most serious weaknesses with the research plan. Often program staff, if they were able to attend the review, can be helpful with this determination. In responding to these major criticisms you should be "appreciative" to the Review Committee for pointing out "weaknesses" in your original proposal and spell out point by point a response to these critical issues. You should devote considerable time to developing a **thoughtful** strategy to address these criticisms either by modifying aspects of the design or analysis, by providing additional information, rationale, or documentation, etc. and then to explicitly address how you would address each concern. In the 3 preliminary pages of the resubmission application you should address briefly how you have addressed the concerns raised in the initial review and point to where in the text of the application the modifications, clarifications, or justifications are discussed in more detail. You should be careful that changes in the design of the study that you make to respond to committee criticism do not adversely affect other aspects of the research (e.g. adding another site to recruit subjects in response to a criticism of not enough subjects for adequate power may create new difficulties for the revised research plan if the second site has a very different population base than the original research site such that the two samples are not comparable on important demographic variables).

In general it is not a good strategy to completely rewrite the application as if it were a new application but incorporating, in general, the committee's criticism. This is because the committee will judge the adequacy of your resubmission by **how well you responded to the specific concerns raised by the review committee in the prior Summary Statement**. If you completely revise the application so the entire project is much improved, but do not adequately address one or more of the specific concerns raised by the committee, they will judge your response to be inadequate. In this regard to ensure adequate space in the application to discuss your response to each of the issues raised by the review committee in the prior review you should economize on the parts of the original application about which the committee did not raise concerns. That is, if the

review committee commented that your literature review was adequate and raised no issues about it, then you can reduce the space you devote to the literature review (i.e., summarize, but not eliminate, the more detailed discussion in the original application) so as to have more space available to address the problematic issues raised in the prior review.

Responses that will doom your resubmission are to ignore the Committee's criticisms, to uncritically change your research design in an attempt to "please" the committee by conforming to perceived committee suggestions, or to denigrate the Committee's criticisms. If you view the criticism as constructive suggestions rather than a personal attack on your research competence, you will find that the suggestions will often significantly strengthen your research. It is often helpful to discuss the Committee's criticisms and your planned revisions with either or both the Scientific Review Administrator of the Review Group and program staff.

Types of Grants

This paper is oriented toward applying for a regular research grant (which we term an "R01"). Such a grant can be for up to 5 years and is theoretically not limited in terms of its research budget (in actuality since grants are funded out of individual branch budgets and not the NIMH budget as a whole, very large project budgets will severely crimp a branch's budget for other areas and is likely to be negotiated down). There are available other types of grant awards; most of which have budget and/or year limitations and/or eligibility restrictions (especially restricted to new investigators). Among these other awards are training awards (both pre- and post-doctoral fellowships), career development awards (termed K awards and targeted at investigators from the beginning stages of their research careers to senior level), B-Start Award (for beginning behavioral researchers and limited to \$25,000 in research costs for a maximum of 1 year), small grants (limited to beginning researchers or established investigators initiating research in a new area and for up to 2 years at a maximum of \$50,000 in research costs per year) and Developmental Grants (R21s) for Psychosocial Treatment Research (limited to preliminary studies of psychosocial treatment approaches for up to 3 years at a maximum of \$100,000 per year). You should consult the specific announcements for each award and program staff about eligibility requirement for these and other awards which might be suitable for you. Grant mechanisms available at the NIMH are described on the NIMH Website at <http://www.nimh.nih.gov/grants/grantgen1.htm#n2>. Announcements for the various mechanisms can be found by scrolling down the list of program announcements for the specific announcement at <http://www.nimh.nih.gov/grants/grantgen1.htm#n2> for research grant mechanisms. Program announcement are listed by year of publication in the NIH Guide so for example the small grant program announcement is PA97-015 and the R21 announcement is PA93-093. There is no announcement for the R01 mechanism instead the general instructions for the PHS 398 grant application provide instructions for submitting an R01 application. In general a regular research grant (R01) is the most competitive to obtain so you should investigate the suitability and your eligibility for these other types of awards. Criteria for evaluating these other types of award differ from

those for an R01. Announcements for research training and career development mechanisms can be accessed at

<http://www.nimh.nih.gov/grants/rtcd.htm>.

Support for New Investigators

Prior to June 1998, NIMH supported a specific type of research project grant, the R29 First Award, that was dedicated to the support of new researchers who were within 5 years of completion of their dissertation or other research training. Beginning June 1998, First Award applications were no longer accepted (see NIH Guide, Volume 26, Number 38, November 21, 1997 and NIH Guide, Volume 26, Number 40, December 19, 1997). A major difficulty with the First Award was that the amount of the award was capped at an average of \$70,000 a year and since it also required a commitment of 50% time of the Principal Investigator there was very little money left over to support the research *per se*. Rather than periodically increasing the First Award budget limits, NIH has opted to channel new investigators into submitting R01 applications (which are not limited in budget and may be up to 5 year for a project) with a procedure to allow identification of new investigators. Review committees would be counseled to consider the new investigator status of the applicant in their review of the application. The exact mechanism to be used to designate the R01 as a project of a new investigator has, at this time, not been finalized. It might be by means of a checkbox on a revised PHS grant application form or a program announcement to be entered on line 2 of the facepage of the application. This new policy would allow reviewers to apply the same (in general, more lenient) criteria for reviewing First Award applications to R01 applications of new investigators without the severe budget limitations of the old First Award. Alternative for new investigators are B-Start Awards (gopher://gopher.nimh.nih.gov:70/00/grants/research//94002 on the Internet or code 940002 through Fax 4U (301-443-5158)) and Small Grants.

Small Grants

Small Grants provide support for up to two years and maximum direct costs of \$50,000 per year. Small grant applications use the same application form as regular research grant applications, but are only allowed 10 pages to describe the research plan. There are eligibility categories for small grants that enhance, but do not preclude, funding prospects.

A good procedure to follow in writing a proposal is to submit a preliminary written version of your application to several colleagues who are knowledgeable in the area of the proposed research and who, preferably, have received a research grant award and/or been on a review committee and ask for them to point out any weaknesses in the rationale, design, or proposed data analysis, so that you can make appropriate revisions before you formally submit the proposal. If you complete a preliminary or final version of your research proposal, you might contact NIMH program staff and ask if they would have an opportunity to provide their impressions of your proposal, but you must realize

that staff impressions of the research will not necessarily be the same as that of a review group.

I hope this information will be of use to you. Further information on the NIH grant review or funding process, can be obtained by contacting NIMH program staff.