

**To Appear in Mental Health, United States, 2003
U.S. Department of Health and Human Services, Substance Abuse and Mental Health
Services Administration, Center for Mental Health Services**

**A BRIEF HISTORY OF EVIDENCE-BASED PRACTICE AND A VISION FOR THE
FUTURE**

**H. Stephen Leff, Ph. D., Director:
The Evaluation Center@Hsri**

October 26, 2002

Evidence based practices is a new metaphor to me. I thought evidence was the source of mental health practices. Are you saying it was not? May still not be? On what are practices then based? At no time have I ever heard that people in the mental health profession based their acts upon anything other than evidence. Was that a lie? Is this "new" metaphor also a lie? And how would I tell the difference? (Personal Communication, Harold A. Maio, MA)

The Evaluation Center@HSRI is a federally funded technical assistance center for the evaluation of adult mental health system change. The Center encourages dialogue with the public about its activities. Recently, the Knowledge Assessment page of our web site received the thought provoking message shown above. The message eloquently states the questions and concerns many people have about evidence-based practices in mental health. This paper addresses the various questions and concerns raised in the message.

First, the paper briefly reviews the history of evidence-based practice. It then describes the types of individuals and organizations currently focusing on evidence-based practices in mental health and nature of the information they provide. Following this, the paper considers the concerns raised about evidence-based practices in mental health. The paper then concludes with a vision of how evidence-based practices should be pursued in the future, taking into account the concerns that have been raised.

In this paper, we focus on psychosocial interventions for adults with severe mental illness treated in the public sector. Nevertheless, we believe our discussion of the concerns about evidence-based practices and our vision for the future are applicable to evidence-based practices in general.

It is important to critically review the concept of evidence-based practices because the move to such practices has become a movement. By a "movement" we mean an organized effort of leaders and followers to identify, disseminate, and cause the adoption of certain practices believed to be different and better than current ones. In summary, the paper speaks to the questions: What are evidence-based practices, where did they come from, why are some people so concerned about them, and where are they going?

WHAT ARE EVIDENCE-BASED PRACTICES?

Evidence-based practices refer to practices that have been tested employing certain scientific methods and shown to be safe (although they may have side effects) and effectiveⁱ for most persons with a disorder or problem. These methods are like, but not necessarily identical to, ones required by the Food and Drug Administration. Below, we will have more to say about these methods. However, there is no FDA for mental health treatments other than drugs, and there are some disputes about how close we can come to scientific standard in testing mental health interventions. For those who believe we cannot truly approximate such standards evidence-based practices are only a metaphor (that is they are the products of processes that have some qualities of science, but are not literally science based).

You can tell whether a practice is evidence-based or not, in the sense we mean it here, if there is a body of scientific research showing that the practice has specific effects that are replicable independent of who does the research. This involves either consulting secondary reviews of studies or synthesizing the results of single studies. We will argue later that ideally, there should be guidelines for identifying evidence-based practices, involving meta-analyses of research findings that quantitatively synthesize the available information.

As Harold Maio's message suggests, many mental health professionals follow practices for which they believe there is evidence. However, the evidence is their own experiences, the experience of their teachers, or the experiences of their clients (this is often referred to as anecdotal evidence). This type of evidence is limited. For one thing, it does not distinguish between changes that happen as a result of treatment and those that happen because of factors such as maturation, the assistance of friends, family, and community caregivers and the passage of time. Secondly, anecdotal evidence may describe the experiences of only a self-selected group of persons, not the experience most people will have. Thirdly, anecdotal evidence is subject to bias. Caregivers and service recipients wish treatments to work for many reasons, some humanitarian, some financial, some ideological etc (Chambless & Ollendick, 2001). The perception that a treatment has worked can be the result of wishful thinking. We are surer that a treatment has worked when independent observers agree that it has. In relying on anecdotal evidence, mental health caregivers are not that different from providers of physical health care. Millenson (1997) cites Eddy as estimating that 85 percent of everyday medical treatments have never been scientifically validated.

Today the mental health field is coming to a different understanding of what is acceptable evidence. This understanding is based on the evolution of medical science and the social and behavioral sciences.

THE HISTORY OF EVIDENCE-BASED PRACTICES

We think of medicine as being based on scientific knowledge. However, if we define scientific knowledge as knowledge derived from true experiments (referred to as randomized clinical trials in medicine) or quasi-experiments that address the threats to validity in other than randomized clinical trials (Campbell and Stanley, 1966), then this has not always been the case. In medicine, analysts point to three landmarks on the road to evidence-based practices.

One is the Flexner Report that created a blueprint for medical education based on a rigorously scientific curriculum (Millenson, 1997). The second is medicine's first randomized clinical trial, a study of the efficacy of streptomycin in treating tuberculosis, which appeared in a 1948 issue of the *British Medical Journal* (Millenson, 1997). The randomized clinical trial placed clinical judgment within a new scientific framework (Millenson, 1977). The third is the establishment of the Food and Drug Administration (FDA) and related governmental organizations with the mission of testing the safety and effectiveness of medical interventions.

Partly due to the publication of Upton Sinclair's novel, *The Jungle*, Congress passed the 1906 Food and Drug Act specifying that the main ingredients of foods and drugs had to be identified on package labels and that the labels could not be misleading (Healy, 1997). The next major event in the evolution of the FDA, as we know it today, was the Food, Drug and Cosmetics Act passed in 1939 which prohibited the marketing of any preparation of a compound until it had been accepted as safe by a newly created agency, the Food and Drug Administration (Healy, 1997). The final defining event in the evolution of the FDA, was the passage of the 1962 Kefauver-Harris drug amendments. This legislation, prompted by birth defects caused by the drug, thalidomide, put in place "more rigorous requirements on manufacturers to satisfy the FDA that a new compound was both safe and efficacious for the ailment for which it was designed" (Healy, 1997, p.26). A major consequence of this act was to institutionalize the view that randomized, placebo-controlled, double blind trials are the gold standard to establish the efficacy of an intervention (Healy, 1997).ⁱⁱ Another major consequence of this act was to underline the

importance of testing interventions for their safety. The FDA does not simply evaluate the efficacy and effectiveness of interventions. It also asks what are the risks associated with this intervention and how do they compare with its benefits. This idea that both the safety and effectiveness of interventions need to be evaluated is another underlying theme in the move to evidence-based practices.

Prior to this, medical practitioners were perceived to have special knowledge, but this knowledge was not necessarily based on science. As Friedson (1970) notes,

...the professional is an expert because he is thought to possess some special knowledge unavailable to laymen who have not gone through his special course of professional training. His special professional knowledge may not be demonstrably and consistently efficacious, but it is the best available to the times, and it is taught to all members of the profession in order to prepare them for the proper performance of their work (p.338).

More specifically, as Healy (1997) notes:

...impressions of both efficacy and safety hinged on the testimonials of a few clinicians rather than on demonstrable effects from multicenter studies and a systematic cataloguing of adverse events (p.26).

Two other advances are important in the evolution of evidence-based practices. The first is the highly influential work on approaches to deriving causal inferences from quasi-experiments by Donald Campbell. Campbell's most well-known work, *Experimental and quasi-experimental designs for research*, appeared in 1966.

The second advance was the emergence of meta-analysis. Meta-analysis is

"a means of combining the numerical results of studies with disparate, even conflicting, research methods and findings; it enables researchers to discover the consistencies in a set of seemingly inconsistent findings and to arrive at conclusions more accurate and credible than those presented in any one of the primary studies. More than that, meta-analysis makes it possible to pinpoint how and why studies come up with different results, and so determine which treatments-circumstances or interventions- are most effective and why they succeed (Hunt, 1997, 1ff).

Meta-analysis was anticipated in early work by Karl Pearson, a British mathematician studying the effectiveness of inoculation against typhoid fever (Hunt, 1997). In 1937, William Cochran, a British biostatistician, developed a key technique of meta-analysis, a method for combining the sizes of effects (effect sizes) reported in different studies (Hunt, 1997). In an influential book published in 1972, Archibald Cochrane, a British epidemiologist drew attention to the fact that people who wanted to make more informed decisions about health care did not have ready access to reliable reviews of the available evidence. Cochrane was highly influential in Britain and the world's preeminent non-governmental organization for conducting research syntheses and meta-analyses, The Cochrane Collaboration, was named after him. We will say more about the Cochrane Collaboration below. However, what Hunt refers to as the "meta-analysis movement" is generally acknowledged to have begun in 1976 with a speech and then a publication by Gene V. Glass (Glass, 1976) of a method for combining studies of psychotherapy.ⁱⁱⁱ

Campbell's contribution was important because he gave coherence and credibility to the idea that, given certain analytic methods, causal inferences could be drawn from studies that

were not strictly experimental. This is important to studying the safety and effectiveness of certain non-pharmacological interventions in health and mental health for which it has proven problematic to conduct true experiments. These problems arise from difficulties in areas such as defining placebo controls, gaining acceptance for random assignment, and controlling the fidelity of practices.

The development of meta-analysis was important to the emergence of evidence-based practices for two reasons. The first is that research and evaluation in science is “anarchic” (Hunt, 1997). It is not organized prospectively into logical steps, except in so far as these are required in research and evaluation activities to meet the requirements of the FDA. Meta-analysis offers a methodology for retrospectively synthesizing uncoordinated studies in a systematic manner that approximating a logical process.

The second reason that meta-analysis was important to the emergence of evidence-based practices is that it offers a route to a relatively unbiased synthesis of the evidence for interventions. Prior to the development of meta-analysis, research was synthesized in narrative review articles (Hunt, 1997) that left substantial room for bias to intrude. As Chalmers and Lau (1993) (quoted in Hunt, 1997) write:

Too often, authors of traditional review articles decide what they would like to establish as the truth either before starting the review process or after reading a few persuasive articles. Then they proceed to defend their conclusions by citing all the evidence they can find. The opportunity for a biased presentation is enormous, and its readers are vulnerable because they have no opportunity to examine the possibilities of biases in the review (p.7).

These advances in education, science, and the role of government have led to a belief in the paradigm of evidence-based practices. Millenson provides a concise statement of this paradigm:

A health care delivery system characterized by idiosyncratic and often ill-informed judgments must be restructured according to evidence-based medical practice, regular assessment of the quality of care and accountability. The alternative is a system that makes life and death treatment decisions base on conflicting anecdotes and calculated appeals to emotion (p.6).

This paradigm has also entered into mental health, as exemplified in the Surgeon General’s Report (U.S. Department of Health and Human Services, 1999). This report states:

As noted throughout this report, a wide variety of community-based services are of proven value for even the most severe mental illnesses... Yet a gap persists in the broad introduction and application of these advances in services delivery to local communities and many people with mental illness are being denied the most up-to-date and advanced forms of treatment... New strategies must be devised to bridge the gap between research and practice.

The remainder of this paper focuses on 3 questions:

The first question is what process should be used to determine the degree to which psychosocial interventions for persons with severe mental illness are evidence-based? Put differently, how can Harold Maio tell the difference between mental health practices that are more evidence-based and ones that are less so? Notice, we believe this question cannot be answered with a simple yes or no. Different amounts of evidence support different psychosocial interventions. Ultimately, we need a system for grading the quality, strength, and consistency of

evidence on a continuum. The history of the evidence-based practices movement suggests the process should involve research synthesis and meta-analysis. However, other factors must also be considered which we discuss below.

The second question is what individuals or organizations should be responsible for supporting this process? For medications, the FDA defines the tests a medication must pass to be accepted as safe and efficacious. Drug companies engage in the arduous task of attempting to meet these tests. Treatments that pass are allowed on the market, ones that fail are prohibited from use. If we wish to know about an available drug treatment, we can find out at least some information about its safety and efficacy from the packaging. Currently, there is no single locus of responsibility for identifying evidence-based psychosocial interventions for persons with severe mental illness, maintaining a registry or database of such practices, disseminating information about the practices, or updating the registry. Instead, multiple actors sporadically and independently support such activities. We consider who these actors are and whether the current situation meets stakeholder needs.

A third question is what are the concerns in identifying evidence-based practices in mental health and how can they be addressed? Will evidence-based practices bring safer and more effective interventions to stakeholders or just provide a new label or “metaphor” for services that have no more foundation in evidence than previous ones? Below, we discuss these concerns and how they might be met.

We turn now to each of these questions.

SOURCES OF EVIDENCE-BASED PRACTICES

In the absence of an FDA for psychosocial mental health practices, various individuals and organizations have assumed the responsibility of identifying those practices that are evidence-based. These include academic researchers, trade organizations, organizations of scientists committed to synthesizing research results, some government agencies with scientific missions, and some advocacy organizations.

A Serious Game With Multiple Players, Multiple Sets Of Rules, No League And No Umpires

The individuals and groups described in this section may not exhaust the players in this serious game. Their activities are well intentioned and have advanced the field. But their activities are also “anarchic” (Hunt, 1997), and therefore confusing to consumers, providers, and other non-scientists. Remaining with the game metaphor, a number of individuals and organizations (players) are behaving according to a loosely specified set and to some extent diverse set of rules. Moreover, there is no “league” in the sense of an organization responsible for defining the rules and assisting people to play by them. And, there are no umpires, individuals who have been given the authority to apply the rules. Therefore we cannot be sure which interventions are winning and which are losing. Also important, we cannot be sure the game is fair, i.e., that all promising practices have had equal opportunity to be tested. These qualifications result in important concerns about evidence-based practices described below.

Unaffiliated Academic Researchers

The lead in identifying evidence-based practices probably has been taken by individual researchers, often in academic settings, who prepare narrative reviews or meta-analyses for

publication in journals and books. Pikoff (1996), for example has compiled summaries and analyses of 242 clinical research reviews published in mental health and substance abuse journals.

These syntheses have no authority beyond that given to them by the reputations of their authors and the journals in which they appear (peer-reviewed journals having the highest credibility) and the scientific quality of the works themselves. Different syntheses adhere to different rules. Some are narrative syntheses, others use meta-analysis. Some include only investigations that were randomized trials, others include quasi-experimental and pre-experimental studies. Over time, syntheses will differ in the studies included. The list of ways in which syntheses can differ is quite long. Not surprisingly, different syntheses can and do reach different conclusions. Nevertheless, there is no accepted process for updating syntheses and resolving differences among them. Nor is there any organization charged with providing the public with updated information from the latest syntheses.

Voluntary Organizations of Scientists Committed to Evidence-Based Practices

Cochran Collaboration

As noted above, Archibald Cochran was a pioneer in the development of methods for synthesizing evidence. In 1992, a group of concerned British scientists at Oxford University established a collaboration to identify evidence-based practices in medicine named in honor of Cochran. The collaboration was formally established in 1993 by 77 individuals from 11 countries. Today the Collaboration consists of 50 collaborative review groups composed of researchers, health care professionals, consumers, and others.

Cochrane reviewers employ methods of synthesizing evidence from work developed by the Cochrane Methods Groups, created to improve the validity and precision of systematic reviews. Currently, methods groups are formed in the following areas: Applicability and Recommendations, Health Economics, Health-Related Quality Of Life, Individual Patient Data Meta-Analyses, Methodology Review Group, Non-randomized Studies, Prospective Meta-Analysis, Reporting Bias Methods Group, Screening and Diagnostic Tests, Statistical Methods.

The purpose of this international body is to aid people to make informed decisions about health care by **“preparing, maintaining and ensuring the accessibility of systematic reviews of the effects of health care interventions.”** The collaboration is based on 10 principles:

- Collaboration
- Building on the enthusiasm of individuals
- Avoiding duplication
- Minimizing bias
- Keeping up to date
- Striving for relevance
- Promoting access
- Ensuring quality
- Continuity
- Enabling wide participation [from website, <http://www.cochrane.org/cochrane/cc-broch.htm#CC>]

Following these principles, the Cochrane Collaboration has created and maintains the Cochrane Library. This library consists of almost 1500 syntheses of medical and behavioral interventions. Using the key words schizophrenia, affective disorder, mental health, and psychosocial interventions we found the 20 entries for psychosocial intervention for adults with serious mental illness shown in Table 2, below. This is by far the largest number of syntheses available from any source. Many more interventions are contained in the library for psychoactive medications and for interventions for children.

The Cochrane Collaboration's principles are ones to which any organization charged with informing the public should adhere. It is striking that the Cochrane Collaboration, a voluntary organization, has accomplished as much as it has. Nevertheless, the Cochrane collaboration's reviews are highly technical and usually involve little input by non-scientist stakeholders, such as consumers and advocates. The Collaboration also makes no effort to reconcile its reviews with those of others that reach different conclusions.

The Campbell Collaboration

In 1999, a concerned group of American scientists founded a collaboration, modeled after the Cochrane Collaboration, to identify evidence-based practices for interventions named in honor of Donald Campbell. The Campbell Collaboration includes researchers from the United States, Great Britain, Canada and Sweden. It is pledged to prepare and maintain systematic reviews of studies of the effects of policies and practices in education and the social and behavioral sciences. Using standards for quality of evidence considered transparent and criticizable, the collaboration has solicited contributions by researchers in fields including criminal justice and substance abuse that meet the needs of those with a strong interest in high quality evidence of "what works." The Campbell Collaboration sees itself as , paving the connection between knowledge assessment and policymaking. The extent to which the Campbell collaboration produces or supports syntheses on psychosocial interventions for persons with serious mental illness, and if it does, how it will coordinate with the Cochrane Collaboration, not to mention other groups, remains to be seen. (Source: <http://campbell.gse.upenn.edu>)

Professional/Trade Organizations

Some professional organizations such as the American Psychiatric Association assess the evidence for practices as a part of guideline development. The National Guideline Clearinghouse (NCG) maintains a searchable database of evidence-based clinical practice guidelines (http://www.guideline.gov/body_home.asp) NCG is sponsored by the Agency for Healthcare Research and Quality, the American Medical Association, and the American Association of Health Plans. NCG can be browsed by disease, intervention and organization. Relevant guidelines can be identified and compared. Guideline comparisons include information on methods to collect evidence, the method used to analyze the evidence, and the methods to assess the quality and strength of evidence. However, almost all organizational guidelines are for disease conditions and not for interventions. These organizational guidelines do not necessarily present the evidence for interventions, although they may provide categories of endorsement for recommendations (e.g., "recommended with substantial clinical confidence" [American Psychiatric Association, http://www.psych.org/clin_res/pg_schizo_process.cfm]). In this paper, we are concerned with organizations that might determine the evidence for practices, not guidelines. So far, the American Psychological Association is the only professional organization that has taken that focus. However, there is a thin line between measuring the

strength of evidence and guideline development as West et al. (2002) note and professional and trade organizations have shown the capacity to assess evidence. The reasons why these organizations might not be the best ones to do this have to do with guild and business considerations evidenced in the experience of the American Psychological Association, described below.

The American Psychological Association

In 1995, the Task Force for Promotion and Dissemination of Psychological Procedures, a task force of the Clinical Psychology Division (Division 12) of the American Psychological Association, issued the first of three reports identifying a number of psychological interventions as empirically supported treatments (Chambless & Ollendick, 2001). The Division 12 Task Force had little to say about the treatment of serious mental illness. It lists only three therapies that “some evidence suggests” are useful in treating schizophrenia and other severe mental illness, family interventions, social skills training, and supported employment). Nevertheless the history and evolution of empirically supported treatments in psychology say much about the problems of expecting provider groups to identify and disseminate evidence-based practices.

The reports of the Division 12 Task Force “reaped both praise and opprobrium” (Chambless & Ollendick, in press). Eventually, the American Psychological Association decided it “would not itself pick up the work of” creating and maintaining the list of evidence supported treatments (Chambless & Ollendick, 2001). These activities are now being continued by a standing committee of Division 12. The Division disseminates information on evidence-based practices through its quarterly publication of the Clinical Psychologist along with a guide to empirically supported treatments in the areas of mental health including Anxiety Disorders and Stress, Depression, Schizophrenia and other Severe Mental Illnesses. Under Depression and Schizophrenia & Other Severe Mental Illnesses, for example, a brief description is given of the disorder along with a narrative summary of psychological intervention with proven results. (URL: http://www.apa.org/divisions/div12/rev_est/index.shtml).

The Division 12 list of empirically supported treatments ran afoul of a number of issues. Among these were “guild or economic” concerns about how managed care might use such a list, fears that practitioners of psychotherapies not on the list would be “disenfranchised”, and the worry that such lists would make practitioners more vulnerable to malpractice suits (Chambless & Ollendick). Any movement to evidence-based practices will have to confront this type of “guild” resistance, when it might be contrary to the interests of providers.

The list of empirically supported treatments ran into technical criticisms, as well (Chambless & Ollendick, 2001). One criticism was that the criteria for deciding what is evidence supported and the methods for reaching decisions were unclear. A related concern was that the criteria were too lenient. Additional concerns were that the available research focused on outcomes that were too narrow and reflective of only certain perspectives. Finally, there was concern that evidence is lacking about whether evidence supported treatments work with subgroups that did not participate in the testing of the treatments and, consequently, also a lack of evidence about how these treatments might have to be modified to be relevant to these subgroups. These criticisms have not only been raised by psychologists about empirically supported treatments. They have been raised by different stakeholder groups about evidence-based practices, generally. We will discuss these concerns in greater detail below.

Division 12 also commissioned A Guide to Treatments That Work edited by Nathan Gorman, and Salkind (1998). This book was produced by a task force of experts separate from the EST task force. A similar publication, What Works for Whom, was prepared by Roth and Fonagy (1996), pursuant to a commission from the National Health Service Executive of the English Department of Health (Roth & Fonagy, 1996)^{iv}. Chambless and Ollendick (2001) note that these different work groups did not use the same categories for indicating the degree to which treatments were evidence supported and they did not define evidence in exactly the same ways. Below, we compare the psychosocial interventions for severe mental illness reviewed by the Cochrane Collaboration, the Division 12 Task Force, Nathan and Gorman, Roth and Fonagy, and Pikoff. Our purpose is to show the unevenness with which these interventions are covered by the different reviewers and to show how findings differ for several interventions reviewed by most sources. However, before doing so, we discuss one more set of players, federal agencies.

Federal Agencies

Although there is a history of government regulation of pharmacotherapies, government agencies as of the writing of this paper, take no equivalent role when it comes to psychosocial interventions for severe mental illness. The theory in this area seems to be that the identification and dissemination of evidence-based practices in mental health can be left to the market place in which consumers armed with information supplied by providers (and occasionally the government) about their choices decide what mental health care they desire. The government does take some responsibility for licensing practitioners, for providing limited funding for research to develop and test new interventions, and for funding promising practices. However, with few exceptions, these efforts are influenced by providers, scientists, and advocates, acting independently and with different information. Consequently, these efforts, although they often make real contributions to care, tend to be reactive and unsystematic, rather than orderly and coordinated.

A completely market-based approach to psychosocial services seems problematic given the evidence that providers can supply biased information and that persons with severe mental illness are a particularly vulnerable group. Below, we discuss several governmental agencies that currently contribute to and support the identification and dissemination of evidence-based practices and whose roles might be enlarged.

The National Institute of Mental Health (NIMH)

NIMH's overall mission is to diminish the burden of mental illness through basic scientific research. It carries out its mission through funding and disseminating research in understanding the nature of mental illness and its treatment, focusing on the areas of basic neuroscience, behavioral science and genetics. The agency consists of five divisions, one of which, The Division of Services Research and Intervention Research houses the Services Research and Clinical Epidemiology Branch. This branch supports and conducts research programs to improve the quality and outcomes of treatment and rehabilitation services. (<http://www.nimh.nih.gov/dsir/index.cfm>). Most of the psychosocial research NIMH funds is field initiated. NIMH takes no particular role in reviewing or synthesizing evidence and it is not its mission to create or maintain an "official" list of evidence-based practices.

The Substance Abuse and Mental Health Services Administration (SAMHSA)

The Substance Abuse and Mental Health Services Administration, located in the Department of Health and Human Services, is the Federal Government's lead agency for improving the quality and availability of substance abuse prevention, addiction treatment, and mental health services in the United States. Its mission is to improve the quality and availability of prevention, treatment and rehabilitation services resulting from substance abuse and mental illness. It accomplishes its mission through exercising leadership by providing strategic funding to increase effectiveness and availability of services. It also develops and promotes quality standards for service delivery, models for training, and effective data collection and evaluation. Mental health services are addressed by the Center for Mental Health Services (CMHS) in SAMHSA. Although periodically CMHS does support the synthesis and dissemination of knowledge to improve mental health services, it has not routinely supported the production, maintenance, and dissemination of this information. This may change as CMHS implements its new 'sciences to services' mission (Curie, 2002). Two concrete manifestations of this agenda are the CMHS supported Evaluation Technical Assistance Center@HSRI (<http://tecathsri.org/>) which maintains a searchable database of narrative, systematic, and meta-analytic reviews of interventions for persons with severe mental illness (EbPMetabase) and the National Association of State Mental Health Directors Research Institute Center of Evidence-based Practices, Performance Measurement, and Quality Improvement (<http://nri.rdmc.org/RationaleEBPCenterReview.pdf>).

The Center for Substance Abuse Prevention (CSAP) within SAMHSA supported the creation of a national registry of effective (drug and alcohol) prevention programs (NREPP) to guide stakeholders in identifying effective prevention programs. NREPP funds an independent contractor to have outside experts in the evaluation of prevention programs review and rate the evidence for the effectiveness of prevention programs. Ratings are made according to 15 criteria including basis in theory, intervention fidelity, process evaluation, sampling strategy and implementation, attrition, outcome measures, missing data, data collection, analysis, possible threats to validity, replication potential, dissemination capability, cultural and age-appropriateness, integrity and utility of the intervention. Prevention programs are solicited to supply the necessary information on a voluntary basis. The NREPP program comes closer than any other government activity to providing, maintaining, and disseminating information on the evidence-base for psychosocial interventions. However, it does not include psychosocial interventions for severe mental illness and its criteria for rating the evidence base for programs do not require rigorous syntheses of all available evidence.

AHRQ Evidence-based Practice Centers

The mission of the Agency for Health Care Research and Quality (AHRQ), formerly the Agency for Health Care Policy and Research (AHCPR), is to support research designed to improve the outcomes and quality of health care, reduce its costs, address patient safety and medical errors, and broaden access to effective services. In the area of identifying and disseminating evidence-based practices for severe mental illness, AHRQ (then AHCPR) joined with NIMH to sponsor development of the Schizophrenia Patient Outcomes Research Team (PORT) Treatment Recommendations (Lehman & Steinwachs, 2001) and the Depression Patient Outcomes Research Team Recommendations (U.S. Department of Health and Human Services, 1999). These recommendations were based on narrative reviews of the schizophrenia treatment

literature. The psychosocial interventions the PORT reviewed were family interventions, vocational rehabilitation, case management and assertive community treatment. The PORT recommendations have been widely disseminated and are currently being updated.

A current AHRQ program that relates directly to the identification and dissemination of evidence-based practices is its Evidence-based Practice program. Under this program, the agency has funded 12 evidence-based practice centers (EPCs) to develop evidence reports and technology assessments on clinical topics using rigorous, comprehensive syntheses and analyses of relevant scientific literature. The EPCs employ meta-analyses and cost analyses to report on clinical topics considered to be common, expensive, and/or are significant for subscribers to federally funded Medicare or Medicaid programs. (source: <http://www.ahrq.gov>).

In 2002, AHRQ released a report on systems to rate the strength of scientific evidence (West, King, Carey, et al., 2002). This report included sections describing systems for grading the strength of bodies of evidence. These systems incorporate judgments of both study quality and whether the same findings have been detected by others using different studies or different people. The report proposes that any system for rating the overall strength of a body of evidence should address three general areas:

- The **quality of findings** measured as the quality of all relevant studies for a given intervention, where quality is defined as the extent to which study design, conduct, and analysis has minimized selection, measurement and confounding biases.
- The **quantity of findings** measured as the magnitude of treatment effect, the number of studies that have evaluated the intervention and the overall sample sizes of the studies considered.
- The **consistency of findings** measured as the extent to which similar findings are reported from work using similar and different study designs.

More specifically, the report proposes that systems for measuring the strength of evidence be rated in terms of the domains and elements shown in Table 1.

Insert table 1 about here

AHRQ would seem to be an agency well suited to identifying, disseminating and maintaining a registry of evidence-based practices. However, its responsibility is all of health care. Psychosocial treatments for severe mental illness are a very small part of all health care. Therefore, it is not surprising that no EPC has conducted any syntheses related to psychosocial treatments for severe mental illness. Nor does the recent report on methods for rating the strength of evidence contain any discussion of the special problems that might be associated with measuring the strength of evidence for psychosocial interventions in mental health.

A Comparison Of Different Reviews

Table 2 compares interventions reviewed by 6 different sources (Cochrane Collaboration, APA, Pikoff, Nathan & Gorman, Roth & Fonagy, Schizophrenia PORT). Of 23 interventions reviewed, only three were reviewed by more than half of the sources.

Insert table 2 about here

Table 3 summarizes the findings by different sources for the three interventions reviewed by more than half of the sources. In this table, we have tried to summarize reviews using the language in the reviews. This table illustrates that except for individual and group psychodynamic psychotherapy (which all reviewers eschew for persons with severe mental illness), reviews present their results using different terms and in some cases reach different conclusions. As noted above, the reviews also used different methods for determining the evidence for interventions. These reviews point to the need for a standard approach to assessing the evidence for psychosocial interventions for persons with severe mental illness, one that adheres to some set of standard criteria (rules) like the ones suggested in the AHRQ report, and that presents conclusions using standardized terms. It also implies the need for an organization to support the identification and maintenance of a registry of evidence-based practices (a league) that presents the latest reviews and decides how to reconcile conflicting conclusions (umpires). As we discuss below, the absence of these and other elements in how we now identify and disseminate information about evidence-based practices has raised a number of concerns about this movement.

Insert table 3 about here

CONCERNS ABOUT THE IDENTIFICATION AND DISSEMINATION OF EVIDENCE-BASED PRACTICES

Below, we discuss six of the most important types of concerns about evidence-based practices. One is that evidence-based practices emphasize the views of scientific elites and funders that pay for their investigations and ignore the views of various citizen groups in a manner that is inappropriate to a democracy that believes in individual participation and self determination (Kitcher, 2001) (the democratic concern). Another concern is that the traditional scientific paradigm for producing evidence-based practices is not the best one for identifying practices that work (the concern that traditional science is limited). Another concern is that there are technical problems in identifying evidence-based psychosocial interventions that have as yet to be resolved. Yet another concern is that for certain subgroups and outcomes they deem important, the evidence for evidence-based practices is slim to non-existent (the overstatement concern). One more concern is that as some interventions are identified as evidence-based, those that are not so identified will be assumed not efficacious or effective (the untested is equal to ineffective concern). A final concern is that establishing that an intervention is evidence-based does not guarantee that the intervention will be implemented with fidelity to the science (the knowing is not practicing concern). Below, we consider each of these concerns and what the response to these concerns has been.

The Democratic Concern

The democratic concern is the idea that consumers of mental health services should participate, but have not participated in research and evaluation about these services. This concern stems from the democratic principle that persons should participate in the decisions that affect their lives. It is a concern that is commonly voiced by consumer advocates and advocates for ethnic minorities (Bernal and Scharron-Del-Rio, 2001; Frese III, et al., 2001; Mazrilli, 2002).

The democratic principle includes a corollary. The corollary is that the values of consumer groups are likely to differ from those of scientists and that these values should influence what interventions are investigated, how risks and desired outcomes are defined, and

other aspects of research methods. Bernal and Scharron-Del-Rio (2001), for example, are invoking the democratic principle and its corollary when they argue that mainstream therapeutic approaches promote individualistic rather than the collectivist or interdependent values which minorities often endorse and that these differences require the development of “methodological pluralism” in testing interventions. Gomory (1999) is invoking the principle and its corollary when he states, “professionally defined expectations of client change can be coercive and patronizing, and ultimately harmful (p.7)”.

This principle has been challenged by those who argue that traditional science is the best method we have for determining what services work and that the views of non-scientists have nothing to do with this. This view has been fostered by the development of evidence-based practice. Healy (1997) states another consequence of the emergence of the FDA was the establishment of the view that medical experts should decide what constitutes a disease and what constitutes outcomes for the treatment of that disease. He also notes that since the founding of the FDA, critics have been troubled by the cooperation between government agencies and manufacturers in setting the rules of the evidence development game (Healy, 1997). Kitcher refers to the situation in which all research decisions are made by “scientific subcommunities” as “internal elitism.” He refers to the situation in which all research decisions are made by scientists and “a privileged group of outsiders, those with funds to support the investigations” as “external elitism” (p.133). Below, we describe a scientific process that blends scientist and stakeholder participation in a manner that Kitcher labels “well-ordered science”.

The Concern That Traditional Science Is Limited

This concern is that the traditional or conventional scientific model is unable to capture the evidence for many practices that providers, ethnic subgroups and consumer groups believe are efficacious and effective interventions. This concern has been around for some time. For example, in an 1835 book, *Research on the Effects of Bloodletting*, Louis wrote that patients who were bled remained sicker for longer and had higher death rates (Millenson, 1997). As Millenson notes:

Not surprisingly, outraged leech users questioned Louis’s methods...[One], for example, warned that mathematical calculations threatened to substitute ‘a uniform, blind and mechanical routine for the action of the spirit and individual genius of the [physician] artist (98ff).

More recently, some psychologists (Chambless & Ollendick, 2001) and “social constructionist” therapists have echoed similar themes (Neimeyer & Raskin, 2001). The most extreme members of this group, “radical-critical” therapists are “dismissive of empirical research” and argue that the performance of their methods requires no justification beyond itself (Neimeyer & Raskin, 2001).

Bernal and Scharron-Del-Rio (2001) provide an example of this concern from representatives of ethnic minority groups. They state that evidence-based practices “developed within the conventional model of science” are of “questionable utility for ethnic minorities, and many of their limitations are a result of or have been rooted in questionable assumptions of the conventional scientific model”. This concern must be addressed by studies of interventions that include adequate numbers of persons from minority groups, that include culturally competent interventions, and address outcomes that members of minority groups desire. However, this does not require abandoning the conventional scientific model.

Concerns About Technical Problems In Identifying Evidence-based Practices

Another set of concerns has to do with technical problems in identifying evidence-based practices. A recent article by Leff & Mulkern (2002) reviews a representative group of these problems. One technical problem is that we need to develop ways of identifying and judging evidence that match the developmental stages of interventions. Below, we will argue that interventions should be developed following a logical progression from clarifying the nature of the interventions to testing them. Different types of evidence will be required for the different developmental stages, however, different guidelines for different interventions at different stages have not been formalized. Another technical problem is that we do not have guidelines for identifying and combining appropriate program contrasts. It is difficult to operationally define placebo-like controls for psychosocial interventions (for example what would a placebo control for supported housing be?). It is also difficult to fully understand the nature of controls described with such terms as “services as usual”. When different studies use different control groups, we will need guidelines for how the data they provide can be synthesized. A third technical problem is that of designing sampling strategies that provide us with evidence about treatment effectiveness, not just efficacy. We cannot synthesize data about how interventions work with various subgroups if the data are not collected. However, collecting data on minority groups is logistically difficult and resource intensive. Psychosocial interventions for persons with severe mental illness, like ACT, include many ingredients. A fourth technical problem is we need methods for synthesizing data to identify the active ingredients of interventions. Otherwise, we run the risk of promoting programs that devote resources to unnecessary or even harmful ingredients. A fifth technical problem stems from the fact that much of the evidence we have about psychosocial interventions for severe mental illness comes from quasi-experimental studies that are analyzed using multilevel models. There is a need for guidelines on how to synthesize data from such models and how to weight such syntheses, as opposed to ones from randomized clinical trials, in making policy and clinical decisions. A final technical problem is that articles in mental health journals do not contain standardized information to facilitate syntheses or to aid replications. Instead, the information is shaped by the preferences of particular authors and reviewers and the pressures on journal space. If there were an FDA-like process that required interventions to pass a standardized series of tests this would not be as much of an issue. However, if syntheses must rest on independent and uncoordinated studies, then the information in journal articles will be crucial. Efforts have been made to improve the reporting of randomized clinical trials in medicine, but such efforts have not been made for psychosocial interventions for severe mental illness (Begg C, Cho M, Eastwood S, et al.; 1996).

The Over-statement Concern

The over-statement concern is that in their zeal to disseminate and inform the public about evidence-based practices, advocates of these practices do not sufficiently stress that there is very little evidence about the efficacy or effectiveness of even our most well-researched interventions for specific subgroups (e.g., ethnic groups) and for specific outcomes (e.g., recovery) (Bernal & Scharron-Del-Rio, 2001; Chambless & Ollendick, 2001; Anthony, 2001). This is a concern shared by representatives of ethnic minority and consumer groups and some scientists. For example, Bernal & Scharron-Del-Rio (2001) state that “from the perspective of the conventional scientific model, we know very little if anything about the efficacy of treatments for ethnic minorities” (p. 333). “Thus, the mission to disseminate and inform the

public should not over-state the...applicability of [evidence-based practices]. [A simple] list [of evidence-based practices] may actually misinform a significant sector of society” (p. 332).

The Concern That Untested Is Equivalent To Ineffective

This concern is primarily a concern of providers who for a variety of reasons prefer to continue to implement practices that are not evidence-based because they have not been subject to scientific testing. Their concern is that payers and potential clients will confuse the fact that their interventions have not been tested with the idea that their interventions are known to be ineffective (Chambless & Ollendick, 2001).

The Concern that Knowing Is Not Practicing

This concern is that even if we identify evidence-based practices, this does not mean that providers will adopt them or implement them with fidelity to the practices that were scientifically tested. As Friedson (1970) and others have noted, there is a lack of “equivalence between knowing and doing”. The result of this gap is a need to develop technologies for motivating and training individuals to implement evidence-based practices with fidelity.

A VISION OF THE FUTURE FOR EVIDENCE-BASED PRACTICE: WELL ORDERED SCIENCE

We believe the concerns raised above can be addressed by a process we will refer to as “well-ordered science”. We have borrowed this term from Kitcher (2001). Kitcher contrasts well-ordered science with “internal and external elitism (p.133).

In well-ordered science applied to determining evidence-based practices, individuals, representing diverse stakeholder groups with different initial values would be brought together by some supporting organization or organizations to discuss the available courses of inquiry. “The product of the consideration [would be] a collection of lists of outcomes the deliberators would like scientific inquiry to promote (and adverse events to be avoided) coupled with some index measuring how intensely they desire those outcomes (p.118)”. Through iterative discussions these preferences would be *tutored* or modified to “absorb the needs” of others with collective preferences reached by consensus or vote. Possible interventional and scientific strategies would then be “*assessed*” to determine how interventions might be developed and tested. These assessments would be used to select a course for developing an intervention for the priority outcomes. The course selected would follow the rules of traditional science, but it could begin by using observational data to identify candidate practices for testing. These practices might not be evidence-based, but would be “best practices”. Best practices are ones that appear best, based on all available information including consumer and provider anecdotes and expert judgment.

The supporting organization or organizations would then fund phased research and evaluation to develop and test the candidate practices (Leff & Hollen., 2002; Leff and Mulkern, 2002). The development phase would consist of writing manuals and workbooks, crafting fidelity measures, selecting or devising outcome measures, and designing training programs (Torrey et al., 200*). The testing phase would address the efficacy and effectiveness of the candidate practices. Testing would involve comparison to no treatment groups or groups receiving placebo or comparison to competing interventions that had been tested. Intervention developers would be required to provide detailed and complete reports of their tests, either in technical reports or publications. Next, the supporting organization or organizations would

convene consensus groups to use systematic methods for assessing the quality, strength and consistency of the evidence for the practices studied (West, King, Carey, et al., 2002). Evidence would be considered from all scientific tests, ideally synthesized by meta-analysis. The group would then give the practice a score or rating. This rating would inform providers and consumers about the probable effectiveness of the intervention. It could be changed by additional research and evaluation. Thus, the supporting organization or organizations would have to support on-going monitoring and be able to convene additional consensus groups. For those interventions that achieved a certain evidence grade, the supporting organization(s) would then fund a dissemination phase in which materials and methods for training providers would be developed to spread the intervention. These materials would include revised and refined manuals, fidelity measures and outcome measures. Finally, the supporting organization(s) would implement a system for post-dissemination monitoring to identify additional uses or risks associated with the interventions discovered in widespread, routine use. This monitoring could also require reconsideration of the evidence for interventions.

Well-ordered science responds to the democratic concern by making a place for consumer and citizen participation in getting to evidence-based practices. The response asserts that science should be primarily about means and while consumer participation should be about ends. “Ends” in the case of evidence-based practices consist of desired outcomes to be sought and undesirable or adverse events to be avoided.

However, well-ordered science does leave us with additional issues about participation. One is the problem of representation (Kitcher, 2001). To implement well-ordered science we will need to operationally define what constitutes adequate representation for stakeholder groups (e.g., how many consumers have to be involved in a consensus group, how should they be selected, and how should their on-going involvement be structured?). Another issue has to do with the time that well-ordered science takes. The more groups that are involved in a process, the longer that process will usually take (Leff & Mulkern, 2002). At the same time, all stakeholder groups are impatient to identify and implement evidence-based practices.

To respond to the concern that the traditional or conventional scientific model is unable to capture the evidence for many practices that ethnic subgroups and consumer groups believe are efficacious and effective, well-ordered science makes a place for the observational evidence that non-scientists find so compelling in the initial phase of the scientific process. To respond to the overstatement concern, well-ordered science envisions organization(s), “tutored” by diverse stakeholder groups convened to play the role of umpire in the testing game, guiding the development and testing of interventions. They decide what should and can be claimed for interventions. To respond to concern that untested interventions are ineffective, the well-ordered science model provides a process accessible to all to have their interventions tested. This still may leave a period of time when providers cannot claim that their interventions have been tested, however, there is no alternative to this short of “grandfathering” certain interventions into the list of evidence-based practices. Given that untested interventions may be ineffective at best, and unsafe at worst, this is not a desirable course for most interventions. To respond to the last concern that knowing is not practicing, the well-ordered science model includes a dissemination phase in which tools are developed to raise the probability that interventions will be implemented with fidelity.

A last and major problem with the vision of well-ordered science is that it does not specify what supporting organization(s) will have responsibility for convening groups, funding

research and evaluation, or deciding when the testing is complete and what its results are. In the development of medications, the pharmaceutical companies fund the research (with some governmental participation) and the FDA attends to the other tasks. As we have seen above, there are no funders with comparably deep pockets to finance the development of psychosocial interventions and there is no individual organization or collaborative that consistently fulfills all of the other roles of the supporting organization(s).

The idea a supporting organization should guide the process of bringing psychosocial interventions to the public is not a recent one. Julian Rotter(1971) proposed this in 1971, and Godfried, 1999) states that, in 1980, Morris Parloff, a National Institute of Mental Health (NIMH) staff member, was called before a congressional committee in which he was asked if NIMH might function much like the U.S. Food and Drug Administration and specify which behavioral interventions were safe and effective. Nevertheless, no one federal organization has been tasked with this role.

As we noted above, some governmental agencies have moved towards being supportive organizations and the model of well-ordered science. However, none have gone far enough to bring about well-ordered science. Given the way in which bureaucratic responsibility is divided, it may be that only some coalition of these organizations can bring about well-ordered science. An example of movement in this direction is the NIMH-SAMHSA Request for Applications for State Implementation of Evidence-Based Practices (EBP)-Bridging Science and Service (RFA MH-030007, 2002). Greater in may be the ultimate challenge posed by evidence-based practices. A process of well-ordered science could address the concerns of Harold Maio and other stakeholders. However, the resources and supportive infrastructure for such a process, not to mention concerns of efficiency, require a coordinated governmental response. Hopefully, one can be forthcoming.

Table 1. Domains and Desirable Elements for Systematic Reviews based on AHRQ Report by RTI/UNC EPC

Domain	Desired Elements^a
Study Question	Clearly specified and appropriate
Search Strategy	<ul style="list-style-type: none"> • <i>Sufficiently comprehensive and rigorous with attention to possible publication biases</i> • <i>Search restrictions justified (e.g., language or country of origin)</i> • Documentation of search terms and databases used provided • Sufficiently detailed to reproduce study
Inclusion and Exclusion Criteria	<ul style="list-style-type: none"> • Selection methods specified and appropriate, with a priori criteria specified if possible
Interventions	<ul style="list-style-type: none"> • Intervention(s) clearly detailed for all study groups
Outcomes	<ul style="list-style-type: none"> • All potentially important harms and benefits considered: symptoms, functioning, service use, satisfaction
Data Extraction ^b	<ul style="list-style-type: none"> • Process is rigorous and consistent • Number and types of reviewers is described • Blinding of reviewers • Measure of agreement or reproducibility are provided • Extraction of clearly defined interventions/exposures and outcomes for all relevant subjects and subgroups
Study Quality and Validity	<ul style="list-style-type: none"> • <i>Assessment method specified and appropriate</i> • Method of incorporation specified and appropriate
Data Synthesis and Analysis	<ul style="list-style-type: none"> • <i>Appropriate use of qualitative and/or quantitative synthesis, with consideration of the robustness of results and heterogeneity issues as a function of subject and study characteristics</i> • Presentation of key primary study elements sufficient for critical appraisal and replication
Results	<ul style="list-style-type: none"> • Quantitative summary statistics and measures of precision, as appropriate
Discussion	Conclusions supported by results with possible biases & limitations taken into consideration
Funding or Sponsorship	<ul style="list-style-type: none"> • <i>Type, sources of support for included studies & review detailed</i>

^aElements appearing in italics are those with an empirical basis Elements appearing in bold are those considered essential to give a system a yes rating for the domain.

^bDomain for which a yes rating is required that a majority of elements be considered.

Table 2. Psychosocial Interventions Reviewed By Six Different Sources

	Intervention	Sources					
		Cochrane	Pikoff	APA	Nathan et al.	Roth & Fonagy	AHRQ Schiz.a PORT
	Schizophrenia						
1	Assertive community treatment	X			X		X
2	Case management for people with severe mental disorders	X					X
3	Cognitive behaviour therapy	X	X			X	
4	Cognitive rehabilitation	X					
5	Community mental health teams (CMHTs)	X					
6	Consumer Operated and Self-Help					X	
7	Crisis intervention for people with severe mental illnesses	X					
8	Day centres for severe mental illness	X					
9	Day hospital versus out-patient care	X					
10	Early Intervention		X				
11	Family Psychosocial (but not Educational) Intervention	X		X		X	
12	Family Psychoeducation	X	X		X	X	X
13	Individual and group psychodynamic psychotherapy and psychoanalysis	X	X			X	X
14	Patient held clinical information for people with psychotic illnesses	X					
15	Prompts to encourage appointment attendance	X					
16	Psychoeducation	X					
17	Seclusion and restraint	X					
18	Short stay hospitalisation	X					
19	Social Skills (life) programs, Social Learning	X	X		X	X	
20	Supported housing for people with severe mental disorders	X		X			
21	Token economy and Behavior Therapy	X			X		
22	Treatment programmes for people with both severe mental illness and substance misuse	X					
23	Vocational rehabilitation (Supported employment [SE], pre-vocational training [PVT], standard care [SC])	X		X			X

Table 3. A Comparison Of Reviews Of Interventions Reviewed By More Than Half Of Sources

Intervention	Sources				
	Cochrane (2002) (Date of study depends on review)	Pikoff, (1996) (Date of study depends on review)	Nathan et al. (1999)	Roth & Fonagy (1996)	Schiz. PORT (2001)
Family Psychoeducation	Possibility of usefulness (2000)	“Moderate success” (1994)	“Several well- controlled studies support use” (1999)	Does not reduce relapse	Reduce one year relapse rate
Individual and group psychodynamic psychotherapy and psychoanalysis	No evidence of any positive effect (2002)	Overall results discouraging (1990)	Psychodynamically oriented therapy not shown to be especially effective	Evidence appears to be largely negative	Lack of evidence makes a strong case against use
Social Skills Training	No clear effect demonstrated (2002)	Gains in specific skills, marginal effects on general functioning and symptoms (1990)	Evidence of helpfulness excellent	Generalized benefit remains to be determined	

References

- Anthony, W. A. (2001). The need for recovery-compatible evidence-based practices. *Mental Health Weekly*, 5(11), 5.
- Begg, C., Cho, M., Eastwood, S., Horton, R., Moher, D., Olkin, I., Pitkin, R., Rennie, D., Schulz, K. F., Simel, D., & Stroup, D. F. (1996). Improving the quality of reporting of randomized controlled trials. *Journal of the American Medical Association*, 276(8), 637-639.
- Bernal, G., & Scharron-Del-Rio, M. R. (2001). Are empirically supported treatments valid for ethnic minorities? : Toward an alternative approach for treatment research. *Cultural Diversity & Ethnic Minority Psychology*, 7(4), 328-342.
- Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research*. Boston: Houghton Mifflin Company.
- Catty, J., Burns, T., & Comas, A. (2002). *Day centres for severe mental illness (Cochrane Review)* (The Cochrane Library, Issue 3). Abstract retrieved July 29, 2002, from <http://www.update-software.com/abstracts/ab0001710.htm>
- Chambless, D. L., & Ollendick, T. H. (2001). Empirically supported psychological interventions: Controversies and evidence. *Annual Review of Psychology*, 52, 685-716.
- Chilvers, R., Macdonald, G. M., & Hayes, A. A. (2002). *Supported housing for people with severe mental disorders (Cochrane Review)* (The Cochrane Library, Issue 3). Abstract retrieved July 29, 2002, from <http://www.update-software.com/abstracts/ab000453.htm>
- Cochrane, A. L. (1972). *Effectiveness and efficiency: Random reflections on health services*. London: Nuffield Provincial Hospitals Trust.
- Cormac, I., Jones, C., & Campbell, C. (2002). *Cognitive behavior therapy for schizophrenia (Cochrane Review)* (The Cochrane Library, Issue 3). Abstract retrieved July 29, 2002, from <http://www.update-software.com/abstracts/ab000524.htm>
- Crowther, R., Marshall, M., Bond, G., & Huxley, P. (2002). *Vocational rehabilitation for people with severe mental illness (Cochrane Review)* (The Cochrane Library, Issue 3). Abstract retrieved July 29, 2002, from <http://www.update-software.com/abstracts/ab003080.htm>
- Curie, C. (2002) Testimony to the House Appropriations Subcommittee on Labor/HHS/Education Hearing. Washington DC.
- Freidson, E. (1970). *Profession of medicine: A study of the sociology of applied knowledge*. New York: Dodd, Mead & Company, Inc.
- Frese III, F. J., Stanley, J., Kress, K., & Vogel-Scibilia, S. (2001). Integrating evidence-based practices and the recovery model. *Psychiatric Services*, 52(11), 1462-1468.

- Gingerich, S. L., & Bellack, A. S. (1995). Research-based family interventions for the treatment of schizophrenia. *The Clinical Psychologist*, 48(1), 24-27.
- Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *The Educational Researcher*, 10, 3-8.
- Godfried, M. "The Pursuit of Consensus in Psychotherapy Research and Practice." *Clinical Psychology: Science and Practice*, 1999, 6 (4), 462-466.
- Gomory, T. (1999). Programs of Assertive Community Treatment (PACT): A critical review. *Ethical Human Sciences and Services*, 1(2), 1-17.
- Hayes, R. L., & McGrath, J. J. (2002). *Cognitive rehabilitation for people with schizophrenia and related conditions (Cochrane Review)* (The Cochrane Library, Issue 3). Abstract retrieved July 29, 2002, from <http://www.update-software.com/abstracts/ab000968.htm>
- Healy, D. (1997). *The anti-depressant era*. Cambridge: Harvard University Press.
- Henderson, C., & Laugharne, R. (2002). *Patient held clinical information for people with psychotic illnesses (Cochrane Review)* (The Cochrane Library, Issue 3). Abstract retrieved July 29, 2002, from <http://www.update-software.com/abstracts/ab001711.htm>
- Hunt, M. (1997). *How science takes stock: The story of meta-analysis*. New York: Russell Sage Foundation.
- Johnstone, P., & Zolese, G. (2002). *Length of hospitalization for people with severe mental illness (Cochrane Review)* (The Cochrane Library, Issue 3). Abstract retrieved July 29, 2002, from <http://www.update-software.com/abstracts/ab000384.htm>
- Joy, C. B., Adams, C. E., & Rice, K. (2002). *Crisis intervention for people with severe mental illnesses (Cochrane Review)* (The Cochrane Library, Issue 3). Abstract retrieved July 29, 2002, from <http://www.update-software.com/abstracts/ab001087.htm>
- Kitcher, P. (2001). *Science, truth, and democracy*. New York: Oxford University Press.
- Leff, H. S., & Hollen, V. (2002, Spring). Talking about translating research into practice. *Outlook*. Cambridge: The Evaluation Center @ HSRI and National Association of State Mental Health Program Directors Research Institute, Inc.
- Leff, H. S., & Mulkern, V. (2002). Lessons learned about science and participation from multisite evaluations. *New Directions for Evaluation*, 94, 89-100.
- Lehman, A., & Steinwachs, D. (2001). *Measuring conformance to treatment guidelines: The example of the schizophrenia PORT*. Cambridge, MA: Evaluation Center At Human Services Research Institute.
- Ley, A., Jeffery, D. P., McLaren, S., & Siegfried, N. (2002). *Treatment programmes for people with both severe mental illness and substance misuse (Cochrane Review)* (The Cochrane

- Library, Issue 3). Abstract retrieved July 29, 2002, from <http://www.update-software.com/abstracts/ab001088.htm>
- Malmberg, L., & Fenton, M. (2002). *Individual psychodynamic psychotherapy and psychoanalysis for schizophrenia and severe mental illness (Cochrane Review)* (The Cochrane Library, Issue 3). Abstract retrieved July 29, 2002, from <http://www.update-software.com/abstracts/ab001360.htm>
- Marshall, M., Crowther, R., Almaraz-Serrano, A. M., & Tyrer, P. (2002). *Day hospital versus out-patient care for psychiatric disorders (Cochrane Review)* (The Cochrane Library, Issue 3). Abstract retrieved July 29, 2002, from <http://www.update-software.com/abstracts/ab003240.htm>
- Marshall, M., Gray, A., Lockwood, A., & Green, R. (2002). *Case management for people with severe mental disorders (Cochrane Review)* (The Cochrane Library, Issue 3). Abstract retrieved July 29, 2002, from <http://www.update-software.com/abstracts/ab000050.htm>
- Marshall, M., & Lockwood, A. (2002). *Assertive community treatment for people with severe mental disorders (Cochrane Review)* (The Cochrane Library, Issue 3). Abstract retrieved July 29, 2002, from <http://www.update-software.com/abstracts/ab001089.htm>
- Marzilli, A. (2002, Winter). Controversy surrounds evidence-based practices. *National Mental Health Consumers' Self-Help Clearinghouse*, 7(1), 5.
- McMonagle, T., & Sultana, A. (2002). *Token economy for schizophrenia (Cochrane Review)* (The Cochrane Library, Issue 3). Abstract retrieved July 29, 2002, from <http://www.update-software.com/abstracts/ab001473.htm>
- Millenson, M. L. (1997). *Demanding medical excellence*. Chicago: The University of Chicago Press.
- Nathan, P. E., Gorman, J. M., & Salkind, N. J. (1999). *Treating mental disorders: A guide to what works*. New York: Oxford University Press.
- Neimeyer, R. A., & Raskin, J. D. (2001). Varieties of constructivism in psychotherapy. In K. S. Dobson (Ed.), *Handbook of Cognitive-Behavioral Therapies*. New York: The Guilford Press.
- Nicol, M. M., Robertson, L., & Connaughton, J. A. (2002). *Life skills programmes for chronic mental illness (Cochrane Review)* (The Cochrane Library, Issue 3). Abstract retrieved July 29, 2002, from <http://www.update-software.com/abstracts/ab000381.htm>
- Pekkala, E., & Merinder L. (2002). *Psychoeducation for schizophrenia (Cochrane Review)* (The Cochrane Library, Issue 3). Abstract retrieved July 29, 2002, from <http://www.update-software.com/abstracts/ab002831.htm>

- Pharoah, F. M., Mari, J. J., & Streiner, D. (2002). *Family intervention for schizophrenia (Cochrane Review)* (The Cochrane Library, Issue 3). Abstract retrieved July 29, 2002, from <http://www.update-software.com/abstracts/ab000088.htm>
- Pikoff, H. B. (1996). *Treatment effectiveness handbook: A reference guide to the key research reviews in mental health and substance abuse*. Buffalo: Data For Decisions.
- Roth, A., & Fonagy, P. (1996). *What works for whom? : A critical review of psychotherapy research*. New York: The Guilford Press.
- Rotter, J. (1971). On the evaluation of methods of intervening in other people's lives. *Clinical Psychologist, Spring*, 1-2.
- Sailas, E., & Fenton, M. (2002). *Seclusion and restraint for people with serious mental illnesses (Cochrane Review)* (The Cochrane Library, Issue 3). Abstract retrieved July 29, 2002, from <http://www.update-software.com/abstracts/ab001163.htm>
- Torrey, W. C., Drake, R. E., Dixon, L., Burns, B. J., Flynn, L., Rush, A. J., Clark, R. E., & Klatzker, D. (2001). Implementing evidence-based practices for persons with severe mental illnesses. *Psychiatric Services*, 52(1), 45-50.
- Tyrer, P., Coid, J., Simmonds, S., Joseph, P., & Marriott, S. (2002). *Community mental health teams (CMHTs) for people with severe mental illnesses and disordered personality (Cochrane Review)* (The Cochrane Library, Issue 3). Abstract retrieved July 29, 2002, from <http://www.update-software.com/abstracts/ab000270.htm>
- U.S. Department of Health and Human Services. (1999) *Mental Health: A Report of the Surgeon General*. Rockville, MD: U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Mental Health Services, National Institute of Health, National Institute of Mental Health.
- West, S., King, V., Carey, T. S., Lohr, K. N., McKoy, N., Sutton, S. F., & Lux, L. (2002). Systems to rate the strength of scientific evidence. *Evidence Report/Technology Assessment No. 47* (Prepared by the Research Triangle Institute-University of North Carolina Evidence-based Practice Center under Contract No. 290-97-0011). (AHRQ Publication No. 02-E016). Rockville, MD: Agency for Healthcare Research and Quality.

ⁱ The distinction between efficacy and effectiveness is an important one for some discussions of evidence-based practices. However, when this distinction is not being explicitly addressed, we will use the single concept of effectiveness to refer to practices that have been found to be either efficacious or efficacious and effective.

ⁱⁱ The 1962 amendments addressed not only prescription, but also over-the-counter drugs. Healy (1997) reports that it was estimated that there "might be up to half a million OTC products on the market...A preliminary investigation of five hundred suggested that anywhere between half and three-quarters were ineffective (p.27). As a result of FDA scrutiny a large number of 'antidementia' drugs and 'antidepressants' vanished (Healy, 1997).

ⁱⁱⁱ Other social scientists like Robert Rosenthal were also working on related techniques, but Glass was the first to receive wide-spread recognition.

^{iv} Chambless and Ollendick (in press) report that efforts related to identifying evidence supported treatments were also undertaken in Canada and Germany.

DRAFT