Published in final edited form as:
Rehabil Psychol. 2008 August ; 53(3): 279–293. doi:10.1037/a0012963.

Evidentiary Pluralism as a Strategy for Research and Evidence-Based Practice in Rehabilitation Psychology

Jalie A. Tucker and
University of Alabama at Birmingham, School of Public Health Birmingham, Alabama, U.S.A

Geoffrey M. Reed
International Union of Psychological Science, Madrid, Spain

Abstract

This paper examines the utility of *evidentiary pluralism*, a research strategy that selects methods in service of content questions, in the context of rehabilitation psychology. Hierarchical views that favor randomized controlled clinical trials (RCTs) over other evidence are discussed, and RCTs are considered as they intersect with issues in the field. RCTs are vital for establishing treatment efficacy, but whether they are uniformly the best evidence to inform practice is critically evaluated. We argue that because treatment is only one of several variables that influence functioning, disability, and participation over time, an expanded set of conceptual and data analytic approaches should be selected in an informed way to support an expanded research agenda that investigates therapeutic and extra-therapeutic influences on rehabilitation processes and outcomes. The benefits of evidentiary pluralism are considered, including helping close the gap between the narrower clinical rehabilitation model and a public health disability model. KEY WORDS: evidence-based practice, evidentiary pluralism, rehabilitation psychology, randomized controlled trials

Over the past two decades, medicine and then public health and the mental health professions have embraced the evidence-based practice (EBP) movement, which seeks to bring health care practices and policies in line with the best scientific knowledge (Brownson, Baker, Leet, & Gillespie, 2003; Goodheart, Kazdin, & Sternberg, 2006; Institute of Medicine [IOM], 2001, 2005; Jenicek, 2003; Norcross, Beutler, & Levant, 2006; Vitora, Habicht, & Bryce, 2004; Westen, Novotny, & Thompson-Brenner, 2004). Rising health care costs and the growth of managed care in the United States in the late 20th century energized the early EBP movement, as did the rapid growth of the scientific knowledge base. The EBP movement is now an international movement that focuses on "evidence" as a guiding force in health care practice and policy. Although the concept of evidence remains at the heart of the movement, the nature and scope of the relevant evidence continue to evolve as EPB has expanded from its origins in acute medical care and drug therapies to encompass care for chronic health conditions and mental and substance use disorders. This paper is about that evolution and how it has affected

---

Correspondence concerning this article should be addressed to Jalie A. Tucker, Department of Health Behavior, School of Public Health, University of Alabama at Birmingham, 1665 University Boulevard, 227 RPHB, Birmingham, AL 35294. E-mail: jtucker@uab.edu.
Jalie A. Tucker, Department of Health Behavior, University of Alabama at Birmingham School of Public Health, Geoffrey M. Reed, International Union of Psychological Science, Madrid, Spain.
Geoffrey M. Reed is now with the Department of Mental Health and Substance Abuse, World Health Organization, Geneva, Switzerland.

Preparation of this manuscript was supported in part by grant R01 AA08972 from the U.S. National Institute on Alcohol Abuse and Alcoholism and grant R21 DA R21-DA021524 from the U.S. National Institute on Drug Abuse. The authors are grateful to the following colleagues for taking the time to communicate with us about the ideas in this article as they apply to rehabilitation psychology (though they should be held blameless as to its content): Susanne Bruyère, John Corrigan, Tim Elliott, Rob Glueckauf, David Keer, Don Lollar, Rhoda Olkin, David Roth, Marcia Scherer, Susan Stark, and Edward Taub.

rehabilitation psychology, sometimes in adverse ways, and why a broadened conception of evidence and the methods used to obtain it may be necessary for advancing the field.

The early EBP movement emphasized the restrictive use of treatments that were “empirically validated” by a minimum of two randomized controlled clinical trials (RCTs) and used the findings as the basis for guidelines and manuals for best practices (Agency for Health Care Policy and Research, 1993; Chambless et al., 1996; Sackett, Strauss, Richardson, Rosenberg, & Haynes, 2000). RCTs came to be viewed as the “gold standard” of evaluation research and as superior evidence compared to non-randomized observational studies of treatment effects and other forms of clinical research (Sacks, Chalmers, & Smith, 1982; Westen et al., 2004). Table 1 shows the well known “evidence hierarchy” (e.g., Jenicek, 2003, p. 34) that resulted from this view of research quality as it informs practice. Strength of evidence is ranked from high to low (Levels I to V) based on the presumed capacity to support causal inferences. Experimental research as epitomized by the RCT is at the top. Later elaborations placed systematic reviews and meta-analyses of RCTs above individual RCTs in the hierarchy (Public Health Information & Data Tutorial, 2006). The emphasis on RCTs contrasts with the broader view of evidence common in public health practice, which values observational and quasi-experimental studies in addition to RCTs (Brownson et al., 2003).

The public health perspective notwithstanding, experimental methods have been increasingly applied to health-related research questions, often without much regard for whether the methods suited the questions at hand. The scope of acceptable methodologies has narrowed, and many scientific journals now require treatment-related research to be reported in line with the *“Consolidated Standards of Reporting Trials,”* or CONSORT, statement (Moher, Schulz, & Altman, 2001), which includes a checklist in which randomization and other typical RCT features figure prominently. The original CONSORT statement was heavily influenced by RCTs for pharmacological treatments and has recently been elaborated to accommodate additional features common to nonpharmacological (e.g., behavioral) treatments (Boutron, Moher, Altman, Schulz, & Ravaud, 2008). Although these methodological directives have helped advance the knowledge base for practice, important research questions that are not amenable to experimentation have been under-resourced or gone unaddressed. Disciplines and fields of inquiry with content questions that are not well suited to experimentation using RCTs are at risk for marginalization, including rehabilitation psychology.

Reed and Eisman (2006) argued that EBP can best be understood in the context of contemporary patterns of organized health care delivery. EBP is a “public idea,” meaning an idea that both describes a public problem—in this case, the undeniable failings in the U.S. health care system—and suggests the wisdom of a particular response—that health professionals should practice in ways that are more consistent with research findings (Tanenbaum, 2005). EBP has been used to justify the lay management of professional behavior, which is the central operating principle of managed care, in ways that reduce health care costs, typically through strategies such as limiting access to services and increasing out-of-pocket consumer costs. Although championed in some quarters (Carpinello, Rosenberg, Stone, Schwager, & Felton, 2002; Chorpita et al., 2002; see Tanenbaum, 2005), there is a dearth of evidence that use of EBP principles by managed care or other health systems results in better and cheaper care (Reed & Eisman, 2006; Westen et al., 2004). There also has been remarkably little challenge to the idea that serving managed care in this way should be a core aspect of the scientific agenda.

Historically, rehabilitation psychology has been allied with and influenced by medical rehabilitation. Both fields have been affected by the growing dominance of EBP and the emphasis it places on RCTs as the most desirable form of evidence. Fuhrer (2003), a former Director of the National Center for Medical Rehabilitation Research (NCMRR), characterized this state of affairs as reflecting “the zealotry of some advocates … who would only admit evidence from RCTs in deciding about a treatment's effectiveness or who, as reviewers of funding applications, insist that the only worthy outcomes research design is the RCT" (p. S11). As a lower status field among medical specialties, rehabilitation medicine has been vulnerable to the charge that its scientific foundation is weak due to a relative paucity of RCTs.

Increasingly, arguments have been advanced for a shift from a sharply hierarchical or vertical view of evidence to a broader horizontal conception that recognizes the limitations of RCTs for addressing some research questions that are central to promoting best practices and policies (e.g., Brown et al., 2003; Concato, Shah, & Horwitz, 2000; Horn, DeJong, Ryser, Veazie, & Teraoka, 2005; Leichsenreng, 2004; Reed & Eisman, 2006; Tucker & Roth, 2006). These arguments endorse the core assumption of EBP that evidence can improve health care practice and policy, but have critically considered what constitutes evidence and how to evaluate its quality and applicability. Tucker and Roth (2006), for example, argued against using fixed "rules of evidence"—such as design, statistical, and effect size requirements—in a content-free fashion. Such features matter in evaluating research quality, but "… can limit flexibility, creativity, and diversity if rigidly applied and thus can work at cross-purposes with knowledge development …. [T]he value of different forms of evidence derives more fundamentally from the developmental state of theory, research, and practice in a given field" (p. 919).

In this paper, we take up these issues as they pertain to rehabilitation psychology and advance an argument for *"evidentiary pluralism*," an approach that entails a more balanced and diverse evidentiary base for practice compared to conventional hierarchies that favor RCTs over all other forms of evidence. The term was introduced by Tucker and Roth (2006) and refers to selecting methodologies based on the needs of a given research question, rather than posing questions within the constraints of methodology. The approach should help advance an interdisciplinary approach to science and practice in rehabilitation psychology and builds on similar analyses in other medical areas (Concato et al., 2000), behavioral medicine (Spring et al., 2005), psychopharmacology (Blacker & Mortimore, 1996), mental health (Leichsenring, 2004; Reed & Eisman, 2006; Westen et al., 2004), substance abuse (Tucker, 1999; Tucker & Roth, 2006), public health (e.g., Victoria et al., 2004), health promotion (World Health Organization [WHO], 1998), and consumer behavior (Seligman, 1995) fields.

There is no doubt that RCTs are invaluable for addressing questions about treatment efficacy. However, the design has often unrecognized limitations that can make it suboptimal for studying other important issues that are intrinsic to psychosocial treatments and behavior change processes and are important for improving the reach and effectiveness of practice. Some questions are better addressed—and sometimes can only be addressed—with non-experimental designs, or require a combination of the two approaches (Brown et al., 2003). This paper challenges readers to evaluate critically and select different research strategies and tactics based on their strengths and weaknesses in relation to specific theoretical or empirical questions.

The article is organized as follows: We first describe the history, advantages, and limitations of the RCT and then consider how these issues intersect with rehabilitation psychology, including shared concerns with other areas and concerns specific to the field's content focus and knowledge development. We illustrate conceptual issues and research questions in rehabilitation psychology that cannot be addressed satisfactorily with RCTs and require alternative methods that often are multivariate, longitudinal, and observational. We conclude that an informed evidentiary pluralism can help advance linkages between rehabilitation science, clinical treatment, and public health practice, and that psychologists should appreciate and help others to learn that not all research questions with practice implications will fit the RCT model.

History, Benefits, and Limitations of Randomized Controlled Clinical Trials

Relative to the natural sciences, the social and behavioral sciences are recent developments of the 20th century and, despite tremendous advances in some areas, the state of knowledge development remains uneven and somewhat limited. This is particularly the case in the behavior change arena where the determinants and mechanisms of change remain poorly understood, even though tremendous pressures exist to develop efficacious treatments for a wide range of health, mental health, and other societal problems that involve behavior. On occasion, experimental research has been premature, unfeasible, or unethical due to the underdeveloped state of knowledge or due to the nature of the phenomena of interest, despite the appeal of experimentation from a causal inferential perspective.

Philosophers of science (e.g., Kuhn, 1970; MacKenzie, 1977; Pepper, 1970) have argued convincingly that approaches to data collection and theory construction should vary depending on the state of knowledge development in a given field. When a science is young, as is the case in many behavioral sciences, descriptive data collected using sound measures are often most useful for advancing knowledge and supporting the inductive development of concepts and theory. As a science matures, a deductive approach to theory construction and hypothesis testing is better supported and facilitative of knowledge development. This generally suggests that a good match between developmental stage and scientific methods will advance a particular field, whereas a mismatch will slow, if not misdirect, knowledge development. But in either case, the scientific method should not be reduced to or confused with mere experimentation, and experimentation is not always possible or desirable (e.g., consider the natural science of astronomy). While experiments often support causal inferences most readily, other methods of data gathering can suggest and illuminate causal relationships. For example, the field of epidemiology, the seminal science of public health, developed through careful observational research on patterns of epidemics that could not be studied experimentally using random assignment. This work resulted in a range of non-experimental methods that have precise rules of inference based on their variable strengths and limitations, and non-experimental methods continue to figure heavily in evidence-based public health practice (Brownson et al., 2003).

Despite these arguments for matching methodologies with the state of knowledge development, applied research areas with public health impact, including rehabilitation psychology, continue to be pushed towards devising the “best” treatment or policy that reduces costs and measurably improves short-term health outcomes. This has understandably resulted in heavy resource allocation for the development of efficacious therapies, and the logic, design, and methods of the RCT have been prominent in this undertaking. RCTs have contributed much to the knowledge base and will remain central to evaluating treatment efficacy, but they are less well suited for investigating other applied questions. This section describes the characteristics, benefits, and limitations of the RCT following the discussion of others (e.g., Horn et al., 2005; Leichsenring, 2004;Rosen, Manor, Engelhard, & Zucker, 2006; Tucker, 1999; Tucker & Roth, 2006) and lays a foundation for the later consideration of longitudinal and observational methods that can be used in place of, or in combination with, RCTs. Such methods are needed to address applied research questions, such as how treatment effects interact with participant and contextual features to influence the variable course of chronic health and behavioral health outcomes.

History of the RCT

RCTs developed in the 1940s to establish the safety and efficacy of new drug therapies. The design was rapidly and widely adopted (Concato et al., 2000), including as part of the U.S. Food and Drug Administration (FDA) process to evaluate new drug therapies for humans after drugs showed promise in pre-clinical trials with animals (Meadows, 2002). The 4-phase FDA process includes small studies with healthy humans (*N* = 20 to 80) to evaluate drug safety (Phase 1); moderate (*N* = several hundred) and then large (*N* = several hundred to several thousand) sample studies to evaluate treatment efficacy for persons with a disease or at risk for a disease before a drug is approved (Phases 2 and 3, respectively); and large studies that extend the populations, dosages, and long-term effects of a drug's therapeutic use after it receives FDA approval (Phase 4). Power is increased across phases to detect both treatment effects and uncommon adverse events that may not be found using smaller samples. The RCT is now widely used to evaluate medical and psychosocial treatments for a wide range of health and behavioral disorders, in addition to pharamcotherapies (Boutron et al., 2008). As discussed next, this extension is based on assumptions that fit well with some research questions, but not with others.

Benefits of the RCT

The RCT is a powerful research tool when applied and executed properly. Much of its appeal lies in the simplicity and benefits of randomization. Random assignment to experimental and control conditions is thought to "neutralize" potential confounding variables, both measured and unmeasured, known and unknown, by distributing their effects evenly across treatment and control groups (Horn & Gassaway, 2007). Potential confounding variables include participant characteristics, self-selection processes, and other contextual features that surround intervention delivery. Following randomization, between-group differences observed shortly after treatment administration support confident inferences that the treatment, and not an associated selection process or confounding contextual factor, caused the observed differences. While maintenance of treatment gains may be clinically desirable, it is not essential for establishing causality.

A cardinal feature of the RCT is its emphasis on maximizing internal validity in order to separate and detect the "true" effect of a treatment independent of other possible influences on outcomes. In addition to the benefits of randomization, the following procedures are often used to aid detection of treatment-produced effects apart from other sources of systematic and error variance: (a) use of large, homogeneous samples with "pure" forms of the condition targeted for treatment; (b) pre-randomization "run-in" periods to select participants who will comply with the treatment protocol; (c) monetary incentives or free treatment for study enrollment or compliance to assure that treated participants are adequately "dosed"; (c) double blind procedures to minimize placebo and expectancy effects; (e) manualized treatment protocols with frequent fidelity checks; (f) standardized outcomes assessment; and (g) wait list or minimal treatment control conditions to assess effects due to the passage of time ("spontaneous" remission).

When used to address an appropriate research question, a well executed RCT provides stronger evidence of a causal relationship between treatment and outcomes than non-experimental research, such as the retrospective case-control or prospective cohort studies that are common in epidemiology. Thus, in the early EBP hierarchy (Table 1), non-experimental designs were placed at lower levels of the hierarchy, both because of their non-experimental status and also because they were thought to entail more risk of selection biases that could affect inferences about treatment efficacy (Sacks et al., 1982). Concerns about selection bias, or the tendency for membership in treatment and control groups to be determined by non-random factors that also influence outcomes, appeared to be valid for some conditions. For example, decades of conventional wisdom and early observational studies had suggested health benefits of hormone replacement therapy (HRT) in post-menopausal women, including a reduction in the risk of stroke. However, a large RCT of HRT efficacy (Wassertheil-Smoller et al., 2003) showed that some forms of HRT significantly increased the risk of ischemic stroke in generally healthy postmenopausal women. Similar discrepancies have been found about the benefits of beta carotene and vitamin A on lung cancer and cardiovascular diseases (Omenn et al., 1996), with observational studies suggesting benefits and experimental studies failing to find them. Such examples are cited as illustrative of the “stainless steel rule of evaluation,” i.e., the more rigorous the research, the lower the chances of positive findings (Rosen et al., 2006).

In spite of these situations in which selection biases appear to have clouded inferences about treatment effects, RCTs and observational research have produced similar results in many other health care areas. Comprehensive meta-analytic reviews have found similar effect sizes for RCTs and well conducted, nonrandomized cohort and case-control studies that used current control groups, selection criteria, and measures similar to those used in the RCTs (Benson & Hartz, 2000; Concato et al., 2000; Shadish, Navarro, Matt, & Phillips, 2000). The findings from the two designs converged across a wide range of treatments and conditions and, when differences were observed, effect sizes tended to be lower, not higher, in the quasi-experimental studies (Shadish et al., 2000). The longstanding concern that observational studies without random assignment will overestimate treatment effects because of selection bias appears to be exaggerated, especially when they employ sound research methods appropriate to that approach.

Limitations of the RCT

A more accurate conclusion about the relative merits of well conducted observational research and RCTs is that both contribute to the evidence base for practice, and confidence about causal inferences and generalizability is high when the findings converge. However, the RCT is not well suited to study how outcomes of interest are influenced by powerful extra-therapeutic variables that may interact with treatment effects over time. As discussed next, this expanded scope of inquiry is essential for understanding rehabilitation processes and outcomes.

Understanding and promoting change in the real world—Although RCTs establish treatments as efficacious under controlled conditions that maximize internal validity, there often are trade-offs between internal and external validity (Shadish, 2002). Concerns about the generalizability of RCT findings to usual practice settings led to increasing emphasis during the 1990s on treatment effectiveness studies as part of the necessary research base for practice (Blacker & Mortimore, 1996; Seligman, 1995; Tucker, 1999). As summarized in Table 2, effectiveness studies evaluate whether treatments established as efficacious in RCTs generalize to real world practice settings and also investigate provider adoption and patient utilization patterns and processes. The study samples better represent the heterogeneity of patient populations and disorders, including co-morbid conditions. Although their use entails greater risk to internal validity, effectiveness studies tend to have greater external validity because they study treatment processes and outcomes in usual care environments using a broader patient mix.

Concerns about how methodological features of the RCT were shaping the knowledge base for practice fueled this expansion of the applied research agenda. Foremost among the concerns is the fact that treatments are never randomly prescribed in the real world (Tucker & Roth, 2006). Complex social, economic, health care access, and health status variables influence service utilization processes and patterns across a wide range of health and behavioral health problems (Anderson, 1995; Mechanic, 1978; Morrisey, 1992; Tucker & King, 1999). But by virtue of random assignment, RCTs create an artificial service delivery environment that eliminates, by design, the role that treatment selection factors have on determining who is likely to enter a treatment program or how treatment selection and other contextual variables may influence treatment outcomes. This trade-off between experimental control and real world conditions is of minor concern when the treatment under study does not depend much on patient motivation, compliance, active behavior change, or the environmental circumstances that surround treatment delivery. But the more these extra-therapeutic dimensions matter in determining outcomes, as they so often do in rehabilitation, the more poorly suited RCTs are for investigating the nature of linkages among the relevant variable sets. For example, RCTs are well suited for studying the effects of relatively passive interventions, such as a surgical procedure that involves little or no post-operative adjustment or rehabilitation by the patient. But RCTs are not a good design choice when the self-selection and motivational components of complying with care matter to its success, as often occurs in a lengthy rehabilitation process. Achieving the desired treatment effect requires more active participation on the part of the patient, for example, by adhering to a home physical therapy program, regularly releasing pressure to prevent pressure sores, properly using and maintaining catheters and other equipment, or not allowing well-meaning family members to offer a level of assistance that interferes with attainment of maximal independence.

It is well established that treatment-seeking and treatment outcomes commonly depend on individuals’ motivation for change; the use of a change process that is consistent with their values, capabilities, and life circumstances; and social network reactions and support (Miller & Rollnick, 2002; Tucker & King, 1999). The RCT, however, treats these person attributes and contextual factors as potential confounding variables, and whatever effect they may have is included in the estimate of the within-group error variance term (Tucker & Roth, 2006). Some studies measure such variables and assess their effects on outcomes statistically, but usually this is done to remove this source of variability from the error term, rather than to understand behavior change and the role of treatment in promoting it. But treatment seeking and treatment engagement are themselves complex processes that are influenced by characteristics of individuals and environments and have important relationships with outcomes (Tucker & King, 1999). Furthermore, the health problems of concern to rehabilitation often wax and wane or follow a deteriorating course that requires multiple treatments and treatment modalities over time (e.g., multiple sclerosis, amyotrophic lateral sclerosis). Understanding the factors that lead to treatment entry at the appropriate point in a disease process in order to benefit maximally from treatment is essential for preventing secondary complications and further disability.

Sample limitations—A second shortcoming of the RCT is the common use of restricted, unrepresentative samples that do not reflect the heterogeneous case mix found in usual care settings or in the larger population with the problem of interest, whether or not health services are sought. Conducting RCTs using carefully selected patients who lack complicating co-morbidities or other contaminating characteristics is not mandated by the RCT methodology. Nevertheless, such homogeneous patient groups are thought to provide the best opportunity to observe treatment effects that are not attenuated by the effects of the multiple contributing factors and other conditions that have been eliminated by the participant eligibility criteria. In rehabilitation, co-morbidities are prevalent (e.g., hypertension, diabetes), and people that have them typically are excluded from RCTs. This may bias effect size estimates of RCTs in the direction of suggesting stronger treatment effects than would be observed in usual practice settings with more heterogeneous and diverse patient groups (Blacker & Mortimore, 1996).

Another concern is the increasingly common use of strict eligibility criteria for sample recruitment, and how these criteria may skew the knowledge base for treatment, sometimes with adverse health outcomes. Horn et al. (2005) discussed such an example involving an influential RCT that demonstrated the benefits of the drug *spironolactone*, a potassium-sparing diuretic, to treat congestive heart failure (CHT) (Pitt et al., 1999). In the 18 months that followed, there was a 4-fold increase in the number of sprionolactone prescriptions, followed by a tripling of hospital admissions and deaths from dangerous elevations of potassium among CHF patients (McMurray & O’Meara, 2004). The problem was that many CHF patients who were prescribed sprionolactone would have been excluded from the RCT that established efficacy because of its stringent selection criteria. In short, the trial's selection criteria did not adequately take into account real-world practice and, in this case, resulted in serious negative outcomes.

Unless interpreted carefully, results of RCTs can be misleading as evidence for practice because of the use of non-random or non-representative samples. The samples are further biased because RCTs only include people willing to participate in a clinical trial that involves random assignment to a promising new treatment or to a usual care or other control condition. This very small proportion of the already small proportion of people with a given problem who will seek and utilize a prescribed treatment may have unusual characteristics. A related concern with relevance to rehabilitation studies is that RCTs are often conducted in university-based tertiary care hospitals that serve those too sick to be treated in their local medical facilities.

Treatment fidelity vs. individualization—A third artificiality of the RCT emanates from procedures aimed at achieving high treatment "fidelity" by standardizing research and treatment experiences across participants. This approach is appropriate when the evaluation target is an acute surgical or pharmacological intervention that would benefit from fine adjustments in technique or dosing. But when the treatments and disorders are distributed through time and are context-dependent, as they often are in rehabilitation and mental health care, rigid treatment standardization risks insensitivity to the dynamic nature and controlling variables of change. It also is difficult to achieve in the clinical rehabilitation context. Most therapists will do whatever they can to help their patients function as well as possible, regardless of whether it is part of a research protocol, and patients may see others receiving beneficial procedures and ask for them, or they may try to replicate them on their own. Such provider and treatment contamination effects may lead to biased estimates of treatment effects in RCTs (Boutron et al., 2008).

Although not well studied, rehabilitation outcomes also are likely influenced by the "common factors" known to influence outcomes in psychotherapy (e.g., therapeutic alliance, patient expectations), which generally are more influential than the technical attributes of different psychotherapies (Norcross, 2002). The ability to form a therapeutic alliance and to motivate patients to participate in lengthy, difficult, tedious, and even painful treatments is at least as important in the rehabilitation context (Purtilo & Haddad, 2007). Such factors cannot be studied in manualized RCTs in which therapists are required to respond in standard ways that are insensitive to variations in client responses. These common research practices promote high treatment fidelity and internal validity, but may contribute to sub-optimal outcomes compared to treatments that are more individualized.

Role of theory—A final concern about RCTs is that the use of random assignment can convey a false impression that there is a reduced need for the guidance of theory in research (Tucker & Roth, 2006). The random assignment feature of the RCT appears to empower atheoretical treatment evaluation research because the investigator does not have to consider, measure, or model the covariates of treatment-seeking, treatment selection, and other extra-therapeutic influences on outcomes over time. This is the case even if the treatment itself is well grounded in theory, and serious missteps can occur in the treatment development and validation process if the mechanisms of therapeutic action are unknown. As Spring et al. (2005) noted, theoretical research about mechanisms of action typically are a preliminary step in the design of biomedical interventions (e.g., based on pre-clinical animal model research). But for psychosocial treatments, concerns about the mechanisms of action typically occur after, not in advance of, RCT demonstrations of treatment efficacy.

In rehabilitation psychology, theory development often has been largely bypassed in favor of evaluating the effectiveness of interventions, products, devices, and environmental modifications. The risks are considerable for advancing the knowledge base, developing interventions, and understanding the mechanisms of intervention action (Fuhrer, 2003; Hollon, 2006). Atheoretical RCTs may indicate that an intervention has an effect, but do not reveal how or why. Furthermore, because many processes important to disability adjustment and participation outcomes (see Elliott & Warren, 2007) cannot be manipulated or subjected to random assignment, theory is especially needed to guide intervention development and the derivation of associated predictions that can be investigated in quasi-experimental research. The disconnection of research in rehabilitation psychology from larger theoretical perspectives on behavior change has meant that most available data are descriptive and do not follow from theoretical propositions that could be used to predict behavior or guide intervention development (Elliott, 1994; Elliott & Frank, 2000).

An excellent example of a theoretically-based program of research that is relevant to rehabilitation psychology grew out of Taylor’s (1983) theory of cognitive adaptation, which proposed that psychological beliefs such as the ability to find meaning in one’s experience, a sense of control, and optimism helped preserve mental health in the face of traumatic or life-threatening (health) events. Although the positive beliefs expressed often were inconsistent with medical evidence and therefore could be considered unrealistic or illusory, their presence was associated with positive adjustment, not with mental health problems (Taylor & Brown, 1988). For example, Taylor’s cognitive theory predicted that unrealistically optimistic HIV-positive men would not be less likely than their more realistic counterparts to engage in safer sex practices, contrary to the opposite prediction of another theory (Weinstein, 1984). A study that tested these divergent predictions supported Taylor’s theory (Taylor et al., 1992).

What makes this line of research powerful and persuasive was the ability to make specific, theoretically-based predictions about what would be observed at each step (see Taylor, Kemeny, Reed, Bower, & Gruenewald, 2000). The prospective studies that comprise this body of work were methodologically rigorous and controlled for many potential confounds, but none were RCTs. This may seem obvious given that no treatment was being evaluated, but our point is that this kind of longitudinal research is better suited for the theory building and evaluation that seem to be among the more significant needs to advance rehabilitation psychology. At a later stage, theoretically-based treatments can be designed and evaluated using RCTs. We return to this issue later with another example involving constraint-induced movement therapy.

Summary

The very features of RCTs that enhance internal validity often preclude the study of the multiple determinants and processes involved in behavior change and that contribute to long term outcomes (Seligman, 1995; Tucker, 1999; Westen et al., 2004). RCTs are an excellent choice when treatments are likely to be efficacious regardless of contextual or self-selection factors and can produce quick benefits that logically can be attributed to the treatment and not to extra-therapeutic events and circumstances that occur during the interval surrounding treatment participation. The design is less suitable for investigating how such contextual variables or treatment selection factors may affect intervention outcomes, or for studying the often lengthy process of behavior change and the role of changing health states in that process. For these reasons, treatment effectiveness studies have developed as a complementary research strategy to the RCT and are aimed at increasing external validity. Interventions established as beneficial in efficacy research are evaluated under usual care conditions using a broader patient mix.

Placing RCTs at the top of the evidence hierarchy tends to drive the development of a knowledge base for practice that favors short-term acute care over chronic disease management and rehabilitation. This is because the kinds of treatments and disorders that can be readily investigated using RCTs tend to be acute medical treatments like surgery and drug therapy for health problems that can be diagnosed and treated in a single, time-limited course of care

(Reed, 2006). The design is not as well suited for studying the long-term functional outcomes of treatments for chronic health and behavioral health problems that have recurring periods of remission and relapse and that are influenced by environmental and social factors, the very conditions that characterize much of rehabilitation. In addition, selecting variables for manipulation, assessment, or statistical control should be a deliberative process guided by theory and prior research, as well as by practical and ethical constraints on what variable domains can and cannot be manipulated experimentally.

Advancing the Applied Knowledge Base for Rehabilitation Psychology

We now return to the question of what kinds of evidence will advance the knowledge base for practice in rehabilitation psychology, in light of the state of development of theory and research. We begin by selectively highlighting the field's accomplishments and then discuss issues in rehabilitation that are difficult to approach with RCT methodologies.

Research Accomplishments that Inform Practice

Rehabilitation has made remarkable accomplishments during its 50-year history, and consistent with the field's multidisciplinary vision, rehabilitation psychologists have been integrally involved in the inception, operation, and direction of the broader medical rehabilitation enterprise and in the development of its research base (Elliott & Frank, 2000). Some of this research figures prominently in a recent IOM report (2007) titled "*The future of disability in America*." For example, psychologists were instrumental in the development and direction of the Model Systems of rehabilitation care for spinal cord, traumatic brain, and severe burn injuries, funded since the 1970s by the National Institute on Disability and Rehabilitation Research (NIDRR) in the U.S. Department of Education. The Model Systems include uniform multi-site data collection and have made substantial contributions to research and practice (IOM, 2007).

These and other initiatives have provided abundant, unequivocal evidence that psychological factors play a powerful role in rehabilitation processes and outcomes. This includes overall outcomes of quality of life and life satisfaction (e.g., Rosenberg, Plakeney, et al., 2006; Tate, Kalpakjian, & Forchheimer, 2002; Whiteneck, Forchheimer, & Krause, 2007) and specific outcomes such as depression, substance use, and suicide risk (e.g., Bombardier, Richards, Krause, Tulsky, & Tate, 2004; Rosenberg, Robert, et al., 2006; Tate, Forchheimer, Krause, Meade, & Bombardier, 2004). Behavioral, psychological, and social factors generally account for more variance in adjustment outcomes than the specific characteristics of an injury or disability (Elliot & Frank, 1996; Elliott & Warren, 2007; Gatchel, Polatin, & Mayer, 1995).

There also is a strong neuropsychological research tradition in rehabilitation, particularly in relation to traumatic brain injury. This work has focused on the relationships among neuropsychological performance, cognitive and neurobehavioral rehabilitation, and outcomes (e.g., Atchison et al., 2004; Niemeier, Kreutzer, & Taylor, 2005), as well as related issues such as the ability to make medical decisions (e.g., Marson et al., 2005). The vocational/occupational research tradition in rehabilitation psychology goes back to the field's inception. This research has produced a rich body of knowledge concerning predictors of vocational outcomes among rehabilitation patients and other people with disabilities and has produced programs that support successful occupational performance, return to work, and job stability (e.g., Walker, Marwitz, Kreutzer, Hart, & Novack., 2006; cf. Elliott & Leung, 2005).

Although it is not always feasible or desirable to dismantle psychological interventions from other components of rehabilitation care, they have generally fared quite well when examined specifically in a rehabilitation context, (Elliott & Jackson, 2005; Elliott & Warren, 2007). Psychological interventions have been shown to make specific and substantial contributions in the context of rehabilitation programs for chronic pain (Fordyce, 1976; Gatchel, 2005), serious mental illness (Corrigan, Mueser, Bond, Drake, & Solomon, 2007), and cardiovascular disease (Linden, Stossel, & Maurice, 1996; Mendes de Leon et al., 2006). Psychologists also have been heavily involved in the development of psychologically-based assessment models to support an optimal match between persons with disabilities and assistive technologies (Scherer, Coombs, Merbitz, & Merbitz, 2006). On a more systemic level, psychologists have contributed extensively to the development, dissemination, and implementation of the WHO (2001) *International Classification of Functioning, Disability, and Health* (ICF) (Reed et al., 2005), including a classification for children and youth (Lollar & Simeonsson, 2005). Despite some acknowledged imperfections, the ICF is emerging as the dominant model of disability (IOM, 2007), and it is increasingly well represented in the rehabilitation psychology literature (Hendershot & Lollar, 2007; Elliott & Warren, 2007).

Unresolved Issues Related to the Poor Fit with RCT Methodologies

In spite of these many accomplishments, psychological research in rehabilitation has frequently been limited by being embedded in larger projects framed by the perspective and research needs of medical rehabilitation. Elliott and Frank (2000) observed that rehabilitation psychologists have often used their research skills in support of overall medical rehabilitation, but, in so doing, have “ignored the need to demonstrate the worth and cost-effectiveness of psychological interventions while nurturing medical databases” (p. 646). They criticized rehabilitation psychology for failing to advance and evaluate psychological models of behavior change, which they say has unwittingly advanced a medical model and has fallen short with respect to providing the evidence demanded in today’s health care environment for psychological interventions. In addition, because much knowledge in rehabilitation psychology does not derive from RCTs, when the evidentiary criteria of the EBP movement are applied, the unsurprising and circular conclusion is that evidence for practice is deficient. We next discuss the ways in which EBP methodological directives have constrained research in rehabilitation psychology and suggest alternative methods for knowledge development.

Shared concerns with other areas of psychological practice—The evidentiary criteria of EBP emphasizing RCTs are more congenial to investigating acute drug and somatic treatments, thereby favoring them over psychological treatments in the evidence base. Rehabilitation psychology and other areas of psychological practice suffer from this bias, which shifts available data in the direction of drug and somatic treatments when evidence is reviewed to determine which practices qualify as evidence-based. Powerful financial incentives also operate to build a research base for pharmacological treatments, which are heavily financed by commercial interests (Reed & Eisman, 2006), and, to a lesser degree, for costly assistive devices and surgical procedures. These incentives include FDA approval, marketplace profits, and substantial hospital billings. This has contributed to the strong bias noted by the IOM (2007) in favor of body-level interventions in rehabilitation research and practice. Comparable financial incentives do not exist for lower-technology aspects of rehabilitation such as psychological, occupational, and physical therapy. As a result, the corresponding research base has grown much more slowly.

Negative findings for somatic treatments are almost never published. This has contributed to further distortion in the evaluation of the research base, because strength of evidence in EBP is generally equated with the number of positive RCTs. Furthermore, many EBP advocates view the double-blind RCT as the ideal clinical trial, including in rehabilitation research (Tate, Findley, Dijkers, Nobunaga, & Karunas, 1999), and some have gone so far as to endorse the randomized blinded trial to eliminate any possibility of investigator bias (Hulley & Cummings, 1988). Obviously, most psychological treatments cannot be studied double-blind, and frequently it is impossible to conduct a single-blind trial.

These factors interact with other preconceptions and biases in the medical community in evaluating evidence for practice, even when research supports psychological treatments. For example, a comparison of cognitive therapy and medication for depressed patients in which the cognitive therapy appeared to have been poorly implemented concluded that its effects were inferior to those of medication (Elkin et al., 1989). But when psychological treatments for depression have been well delivered, they have compared favorably to medications (Hollon, Thase, & Markowitz, 2002; Hollon, Jarrett et al., 2005) and have had longer lasting effects (Hollon, DeRubeis et al., 2005). Nonetheless, influential guidelines for depression treatment, primarily developed by psychiatrists, emphasized pharmacotherapy over psychological treatment (e.g., National Institute for Clinical Excellence, 2004).

Although there are several RCTs of post-acute rehabilitation services as practiced in usual treatment environments (e.g., Langhorne & Duncan, 2001), rehabilitation as an enterprise shares many of the characteristics that make psychotherapy difficult to study using RCTs (Seligman, 1995). First, rehabilitation is generally *not of fixed duration*. In RCTs, the treatment typically stops after a certain interval or number of sessions, regardless of how the person is doing. Second, rehabilitation is *self-correcting*. If one technique is not working, another technique or modality is tried. In RCTs, treatment is confined to specified techniques, administered in a standardized manner. Third, in rehabilitation, patients are generally *active*. They participate in the selection of treatments and goals, which is considered desirable, if not necessary for treatment success. This contrasts with the passive process of random assignment and acquiescence to whatever treatment is offered as part of an RCT. Fourth, individuals undergoing rehabilitation typically have *multiple problems*, and treatment is geared toward relieving parallel and interacting difficulties. Patients in RCTs are typically selected to have a single problem, and co-morbid conditions are often a basis for exclusion. Finally, rehabilitation is concerned with the improvement of *general functioning* as well as the amelioration of specific presenting symptoms, which are generally the focus of RCTs.

Special concerns for rehabilitation psychology—Further characteristics of rehabilitation are difficult to study using RCTs. First, the time horizons for many of rehabilitation's intended effects are extremely long. RCTs typically involve short follow-up periods (e.g., 3–6 months), and inferences about efficacy are based on changes observed during or soon after treatment administration. When patients are followed for longer periods that more fully capture the intended goals of rehabilitation, it is increasingly difficult to attribute observed differences in outcomes to the experimental manipulation rather than to other environmental factors that cannot be controlled. In addition to inevitable changes in patients' life contexts, different phases of the rehabilitation process and recovery trajectory typically involve different needs and goals, and individuals with disabilities may seek and receive many treatments and services over many years. Because of the problems that long time horizons and environmental interactions pose for RCTs of a single acute episode of care, both procedurally and with respect to causal inferences, participation outside of the treatment context is almost never used as a measure of rehabilitation outcomes. Yet it is nearly universally endorsed by consumers, health professionals, and funding agencies as the ultimate goal of rehabilitation (Heinemann, 2005).

One way to deal with this state of affairs is to enroll a large sample that is selected based on the functional rather than physical nature of disability and then follow them for many years using a core assessment battery that is supplemented intermittently by measures tied to the expected trajectory of recovery and that include global endpoint measures of participation. For example, Whiteneck and colleagues (2004) conducted a large longitudinal study of 2,726 individuals with spinal cord injuries that investigated the role for environmental factors in long-term adjustment assessed 1, 5, 10, 15, 20, and 25 years after initial injury. Environmental factors were more strongly related to life satisfaction than to societal participation. Such longitudinal cohort studies could be supplemented intermittently with brief randomized trials of short-term interventions with expected time-limited effects that are based on normative or individual recovery patterns and rehabilitation needs. Although lingering exposure effects from one such treatment to the next would be a concern, treatment exposure would be deliberately sequenced and the effects measured, rather than occurring in an uncontrolled, but probably non-random fashion as patients seek out additional services during the lengthy recovery process.

Second, rehabilitation research often encounters difficulties obtaining sufficiently large, homogenous samples to adequately power RCTs when eligibility is based on diagnosis, because base rates of many conditions are low. As discussed later, this issue may be resolved somewhat by adopting a “disability model” of rehabilitation that focuses more on functional capacity than on the nature of the physical diagnosis, injury, or avenue into the health care system. Use of the ICF (WHO, 2001) to define samples based on functional status rather than diagnosis offers a potential means to obtain sufficient sample sizes for RCTs.

Third, defining and selecting an appropriate control group in RCTs can be challenging in rehabilitation studies. The ethical doctrine of *clinical equipoise* in RCTs requires that (a) there is evidence-based disagreement or uncertainty about best practices; and (b) no participant is “randomized to an intervention known to be inferior to one of the treatments under investigation or the established, scientifically validated standards of care” (Miller & Brody, 2007, p. 153). This is a complex determination and one that has come to favor RCTs that compare a “usual treatment” with an experimental treatment; wait list, minimal, or no treatment control groups have fallen into disfavor on ethical grounds. However, clinical equipoise is particularly is difficult to achieve in rehabilitation RCTs because there often is no “usual treatment” against which an experimental treatment can be compared. Natural history data are generally lacking, especially concerning initial trauma and injury (e.g., stroke, spinal cord injury), so the time frames of natural recovery and intervening markers of longer term adjustment and outcomes that could serve as benchmarks against which to assess further benefits due to rehabilitation are unknown. When coupled with the problems of defining control groups, the lack of natural history data makes it is difficult to conduct treatment efficacy studies in rehabilitation psychology that fulfill the methodological and ethical requirements of sound RCTs.

Collectively, these methodological issues lead in a direction that favors acute medical and pharmacological treatments with short-term outcomes and biological endpoints, and thus may be particularly troubling for rehabilitation psychology. The IOM (2007) characterized research on rehabilitation and disability as “highly skewed toward basic and clinical research, with inadequate support for research on the physical, social, and other environmental contributors to disability and insufficient attention to the evaluation of interventions to minimize activity limitations and participation restrictions” (p. 5).

Conceptual issues—In addition to the constraining influence of RCT methodologies, two unresolved conceptual issues appear to be limiting the advance of the field. The first concerns the diverging research agendas of what we refer to as the dominant clinical *rehabilitation model* and an emerging public health-oriented *disability model*. The clinical rehabilitation model defines its subject matter according to who enters specific systems for rehabilitation care and views the overarching purpose of rehabilitation research to improve systems of care and patient outcomes (e.g., cost-effectiveness, quality of life, health status, physical functioning). Evaluation of specific interventions using RCTs is wholly consistent with this view, even though carrying out such studies may be challenging because of the multifactorial nature of the interventions and the temporally distant nature of many desired rehabilitation outcomes.

In contrast, the disability model defines the target population using an explicit set of criteria based on the associated function and goals (e.g., eligibility for health benefits, access to public facilities). Because the definition of disability depends on its function, no definition is universally accepted (Lollar & Crews, 2003). Disability and morbidity are not synonymous (Chamie, 1995), and people identified as disabled may not consider themselves disabled or experience functional limitations or health problems. The disability model underlies the 2007 IOM report and is closely aligned with a public health model aimed at preventing *secondary conditions* (Lollar, 1999), which is any condition to which a person with a primary disability is more susceptible (e.g., medical, mental, family problems). Preventing secondary conditions is part of the *Healthy People 2010* agenda (Department of Health and Human Services, 2000).

While the disability model research agenda includes a role for the systematic evaluation of interventions, it directs attention toward new research questions, such as predictors of positive adjustment among people with disabilities and the nature of the adjustment process (e.g., Dunn & Dougherty, 2005; Elliott, Kurylo, & Rivera, 2002), reducing employment barriers to people with disabilities (e.g., Bruyère, in press), understanding how people with disabilities experience their identity and treatment in society (e.g., Olkin & Pledger, 2003), and addressing the fact that many common psychology measures have not been normed on people with disabilities and may require significant adaptation (e.g., Olkin, 2002). Its expanded view of interventions includes population-based public health interventions and environmental modifications, in addition to individual treatments. Research guided by a disability model is more likely to be based on functional than diagnostic status, which is important because specific diagnoses are not good predictors of secondary conditions (Lollar, 1997; Ravesloot, Seekins, & Walsh, 1997). Interventions are likely to be seen in the context of an unfolding process rather than as isolated treatments with immediate outcomes. The research agenda that proceeds from the disability model extends beyond treatment efficacy studies that are the foundation of a narrow view of the evidentiary requirements for EBP, and it thus seems better suited than the rehabilitation model to guide much of contemporary research in rehabilitation psychology.

A final issue concerns outcome criteria in rehabilitation research. Full participation in society has long been viewed as the ultimate goal of rehabilitation (Heinemann, 2005; IOM, 2007; Pope & Tarlov, 1991; Wright, 1980; WHO, 2001), including as part of the Americans with Disabilities Act (1990). Although participation is differentiated from activity performance alone and is distinguishable empirically (Cardol, De Jong, & Ward, 2002; Van Brakel et al., 2006), far more attention has gone into developing measures of body functions and activity performance in rehabilitation contexts than into measures of societal participation (Heinemann, 2005). It is much easier to measure performance of daily tasks such as drinking, eating, bathing, and walking than to measure the longer term and more global outcomes that tend to be most important to people with disabilities or who are undergoing rehabilitation following an injury or in the context of a chronic illness, such as: Will I be able to work? Will I be able to have close friendships? Can I manage living on my own? Can I have an intimate relationship? Can I be a good father? These outcomes are of more interest to society as well. However, significant method development work remains before rehabilitation research can focus on outcomes related to participation (Heinemann, 2005). The development of such measures is complicated by the fact that participation outcomes often are temporally distant from specific rehabilitation interventions. Yet, this effort is essential if research is to address the outcomes that individuals with disability say are most important to their lives.

Revising the Evidence Hierarchy to Promote Knowledge for Practice

The preceding discussion highlights conceptual issues and empirical questions that cannot be fully addressed using RCTs. This section summarizes recent expansions of the evidence base for practice in ways that extend its reach and effectiveness in real world settings. Alternative research designs are described that investigate naturalistic components of care-seeking and behavior change either separately or in conjunction with experimentally manipulated treatment variables. Such research complements the evidence base for practice provided by treatment efficacy studies and leads to an expansion of the dimensions of research that are relevant to evaluating the utility of evidence for informing practice.

Revisions to Study Natural Variation and to Model Treatment Selection

Several recent critiques have placed greater value on studies of treatment-seeking and treatment effectiveness in the natural environment (Horn, 2006; Horn et al., 2005; Leichsenring, 2004; Tucker & Roth, 2006; Seligman, 1995; Shadish, 2002). Well conducted naturalistic studies are considered an undervalued source of evidence that are better than RCTs for addressing certain research questions that are vital to practice. As Horn et al. (2005) noted, such methodological alternatives "… do not replace RCTs, but rather provide additional sources of systematic outcomes information …. that answer questions in the real world, where multiple variables and factors can affect the outcomes" (p. S14). "Unlike RCTs, [such] studies do not experimentally alter the treatment regimen to evaluate the outcomes associated with a particular intervention, but take advantage of natural variation in patients and care processes occurring in routine practice …" (Horn, 2006, p. 2732). Leichsenring (2004) similarly emphasized that RCTs of treatment efficacy and naturalistic studies of treatment effectiveness serve different purposes and thus require different criteria to judge research quality. Noting that "psychotherapy is not a drug that … works equally under different conditions" (p. 144), he presented an evidence scheme that placed highest value on prospective quasi-experimental studies with high clinical representativeness that used treatments, patients, and therapists common in practice.

Tucker and Roth (2006), in the context of research on treatments for substance use disorders, specifically emphasized the value of multivariate longitudinal studies of treatment, person, and contextual variables that influence outcomes. The fundamental goal is to model the multiple influences on health and behavior, including but not limited to interventions, and to make valid inferences that apply to larger populations. The multivariate longitudinal approach overlaps with and extends beyond the RCT. Such studies can accommodate experimentally-manipulated independent variables and naturally occurring change in variables that are not manipulated (Singer & Willett, 2003). Specific multivariate methods such as mediation modeling can be applied to both experimental investigations (Shrout & Bolger, 2002) and longitudinal designs (Judd, Kenny, & McClelland, 2001).

Unlike in RCTs, in the multivariate longitudinal model, randomization is not used to avoid the important theoretical work required for model building, even though randomization may be useful for providing targeted experimental control of a key predictor variable (Tucker & Roth, 2006). In a well developed longitudinal model, randomization decisions are based on theory and research in the field and on the ease with which a variable is amenable to controlled manipulation. Randomization can be used to strengthen select causal pathways in an overall multivariate longitudinal model that includes experimentally manipulated and naturally occurring pathways. But most associations will be among the paths that are naturally occurring without experimental control. Inferences to clinical populations thus may actually be strengthened by avoiding unnecessary experimental control, a point echoed by Horn (2006).

Use of multivariate longitudinal models is relevant to rehabilitation psychology because treatment is only one variable of many that contribute to longer-term outcomes. The approach puts treatment in context, allows for the study of participation as a longer-term outcome, and is compatible with the disability model that focuses on the person in context, not just on the person as a patient in the health care system. Treatment can be investigated as an independent or dependent variable. The latter focus is an important complement to treatment efficacy research because treatment seeking, engagement, and outcomes are complex, inter-related processes that multi-determined (e.g., by patient motivation, social support, and health care access, quality, and monetary cost). RCTs, however, preclude the study of treatment selection factors and how they may interact with treatment and other contextual factors to produce long-term outcomes.

Finally, Borckardt and colleagues (Borckardt, Nash, Murphy, Moore, Shaw, & O'Neil, 2008), following Kazdin (1982) and others, argued anew for the direct relevance to clinical practice of practitioner-generated case-based time series designs with baseline measurement. They contended that this design "qualifies as a true experiment and that it ought to stand alongside more common groups designs (e.g., the … RCT) as a viable approach to expanding our knowledge" (p. 77). Relegating case studies *in toto* to the lowest level of the evidence hierarchy misses important distinctions between sound and unsound single-case designs.

Dimensions for Evaluating the Utility of Evidence for Informing Practice

In their extension of the evidence hierarchy to include multivariate longitudinal studies of treatment, person, and contextual variables that influence outcomes, Tucker and Roth (2006) identified five dimensions of research to guide judgments about evidentiary value. The first dimension is whether the aim of the study is to obtain descriptive data to generate hypotheses for future research, or whether the study is designed to test hypotheses and to infer from sample data estimates of effects in the larger population. Although often undervalued, the descriptive or exploratory functions of research are very important in the early stages of knowledge development in a given field (MacKenzie, 1977).

A second dimension is whether the research is cross-sectional or longitudinal. Three or more waves of data are necessary to estimate both true change trajectories and measurement error components separately for each individual participant. Simple two-wave studies with one baseline and one post-treatment assessment generally do not qualify as longitudinal designs (Singer & Willett, 2003).

The third dimension is whether the study analyses and statistical models are bivariate or multivariate. "Multivariate" refers to an analysis with more than one predictor or independent variables (as in multi-determined outcomes) or more than one outcome being modeled simultaneously. Multivariate longitudinal designs are suitable for evaluating mediation and moderation as mechanisms by which treatments have causal effects (Kazdin, 2007;Kraemer, Stice, Kazdin, Offord, & Kupfer, 2001). In contrast, simple bivariate analyses such correlation coefficients or one-way analyses of variance can provide little information about multiple determinants of behavior or about the independence or codependence of multiple outcomes.

The fourth dimension is whether the research goal is to model naturally occurring variability or to test the effects of experimentally manipulated variability. The RCT is the primary method used for the latter purpose, but most behavioral scientists wish eventually to understand and explain phenomena in the natural environment. This dimension directly affects and informs clinical practice. The controlled experiment contributes to this overall goal, but is not the final step in an expanded conception of evidence.

A final dimension of research is its potential public health significance. Public health approaches accept that behavioral interventions tend to have positive, but modest effect sizes and that outcomes are multi-determined. Instead of seeking to maximize treatment benefits for individuals ("clinical significance"), public health approaches seek to reach more of the affected population with beneficial services. Because population impact is a function of intervention effect size and the reach of the intervention into the population in need (Abrams & Emmons, 1997), a less powerful intervention that is used by many can have greater impact on population health than a more powerful and costly intervention received by relatively few.

These are not new ideas, but they have been slow to influence the applied research agenda in rehabilitation psychology and other areas of psychological practice. A population perspective highlights the need for research to increase access to services and to study treatment utilization as a dependent variable (as in health economics) and consumer preferences for services, in addition to studying treatment as an independent variable (the efficacy agenda). Consumer input from persons with disabilities is under-developed in the design and study of rehabilitation services (Olkin & Pledger, 2003) and deserves greater attention.

Revisiting the Purpose of RCTs: Establishing Causality or Standards of Care?

If one entertains, as we have done, that the RCT is not always the most informative evidence for practice, then interesting new questions about its scientific utility emerge, which we take up in this final section. In defending RCTs as the best form of evidence for effective practice, Hollon (2006) stated, “if you can get a pig to fly then you do not need a control group (or statistics) to tell you that you have had an effect. Pigs do not fly of their own accord and, if you can overcome the laws of God and nature, then you have done something” (p. 98). Those who style themselves to be the most die-hard of empiricists might disagree and point out that replicability of the effect and ruling out alternative causes would be necessary to reach conclusions about the treatment, which would best be addressed using controlled research designs. What is more relevant to the current discussion, however, is that Hollon’s proposition reflects a common view in psychology that the primary purpose of RCTs is to establish causality.

In the context of much medical research, RCTs frequently do not have the establishment of causality as their primary purpose (Spring et al., 2005). The rise of RCTs as the gold standard of evaluation research was based on its use to establish the safety and efficacy of drug therapies as part of the FDA process (Concato et al., 2000). By the time a Phase 1 clinical trial begins, basic research to establish the mechanism of a drug’s effect has been finished. The issue, then, is whether the causal effect observed in pre-clinical animal research will translate into therapeutic efficacy in humans with specific clinical conditions. The research questions concern drug dosage and the range of clinical presentations for which the drug is effective.

From a medical point of view, the main purpose of RCTs is to establish standards of care for which health care providers are professionally and legally accountable (Reed & Eisman, 2006). In establishing practice standards, while teaching a pig to fly might indeed be a noteworthy accomplishment, RCTs would still be necessary to establish that the process was safe for the pig, which training techniques and dosages were required to achieve the effect, for what sort of pigs the training worked, and, eventually, the long-term effects of flying on pigs. This has generally not been the psychological approach to RCTs. Instead, RCTs are often expected to carry the burden both of demonstrating therapeutic efficacy and of elucidating the mechanism of the effect, e.g., via the design of a control condition that contains everything but the putative “active ingredient.” Questions about dosing and the range of clinical applicability are addressed in a very general sense in effectiveness studies, but are rarely the focus of systematic inquiry.

Constraint-induced movement therapy (CIMT) is an example in rehabilitation psychology in which use of RCTs to establish standards of care is highly appropriate, but only as one phase in a broader program of research. CIMT was based on neuroscience research with adult monkeys with surgical somatosensory deafferentiation that demonstrated neural plasticity beyond what had been thought possible in adult mammals (Taub, 1980; Taub, Uswatte, & Elbert, 2002). Taub’s animal model on which CIMT is based can be seen as parallel, if not superior, to the information that would be available about the potential therapeutic mechanism of a new drug prior to Phase 1 FDA testing. On this basis, CIMT was developed as a new physical rehabilitation program for persons with extremity mobility problems due to central nervous system damage from causes such as brain injury or stroke (Taub, Uswatte, Mark, & Morris, 2006). CIMT has three components: (a) daily repetitive, task-oriented training of the impaired extremity or function using shaping principles for two to three weeks; (b) constraining patients so they have to use the impaired extremity or function during waking hours; and (c) a "package" of behavioral methods for transferring therapeutic gains to the real world.

CIMT has been found to substantially improve upper limb mobility deficits in patients with mild to moderately severe chronic stroke (Taub et al., 1993). At the time conventional wisdom was that stroke patients would show little improvement in motor function six months after stroke onset, and certainly not after one year, regardless of the therapy administered. Remaining motor deficits were considered permanent, and many patients are still told that. The mean chronicity of Taub et al.'s (1993) patients was over four years. In rehabilitation, therefore, this is the logical equivalent of teaching a pig to fly.

Evidence for the efficacy of CIMT continues to grow, including RCTs and replications by other researchers (Taub, Uswatte, King et al., 2006; Wolf et al., 2006). Research on CIMT over the next decade will likely to continue to rely on RCTs, but not because causality is at issue. Clear evidence of causality comes from the large body of pre-clinical animal research and is buttressed by extensive clinical experience with chronic stroke and other conditions that are the focus of CIMT. The primary goal, then, of the RCT research will be to establish standards of care (see Taub & Uswatte, 2006), addressing such questions as: What is the necessary duration and intensity of CIMT? Are all three components necessary? With what range of patients and conditions is CIMT effective? Additional questions that cannot be addressed by RCTs alone include: How cost-effective is CIMT? What type of training is necessary to perform CIMT? What patient characteristics moderate the effects of CIMT?

Addressing these questions, especially those concerning cost, are necessary (though perhaps not sufficient) to persuade health insurance plans to cover CIMT, which can be seen as an expensive, intensive service for a large group of people who previously had been classified as unable to benefit from further treatment (Taub & Uswatte, 2006). Denials of coverage have been ostensibly based on the CIMT's "experimental" status, even though it enjoys a stronger evidence base than exists for many other covered practices.

Conclusions

As a part of graduate training, psychologists learn research methodology based on principles of logical reasoning and critical thinking. This is a tradition we value enormously. But when the concern is the development of research on clinical interventions, there may be considerable liability to the psychological view of RCTs as being devoted to establishing causality and elucidating mechanisms, which are often better addressed by other research approaches. Although weighing different strategies with which to investigate a given research question imposes a significant scholarly burden and risks "evidentiary dilution" in the hands of uncritical researchers, psychologists should recognize, and help others to learn, that not all research questions—and not all forms of treatment—will fit the RCT model.

The field of rehabilitation psychology seems particularly ill served by overly restrictive methodologies. As discussed earlier, this is due to the field's developmental status and the complex nature of questions the field must address to be of value to patients in rehabilitation care, to individuals with disabilities, and to society. Clearly, the RCT has a central role in outcome research related to treatments in which rehabilitation psychologists participate. However, we believe that it is imperative that rehabilitation psychologists resist following the dictates of a narrow construction of the EBP movement that has elevated research using the RCT and the standardized treatment manual above all else (Tucker & Roth, 2006). In this respect, we concur with a meta-analysis of methodologically diverse studies of medical treatments (Concato et al., 2000) that concluded: "The popular belief that only randomized, controlled trials produce trustworthy results and that all observational studies are misleading does a disservice to patient care, clinical investigation, and the education of health care professionals" (p. 1892).

We contend that positive developments will come from an informed and critical evidentiary pluralism, and that this will help promote the development of a interdisciplinary research base for rehabilitation that includes psychology in a leadership role. The efficacy-effectiveness distinction highlights how different empirical questions are better addressed using different methods and how, when considered together, can expand the scope and impact of health-related research. If both are components of the field's research portfolio, the trade-offs between internal and external validity are more balanced in the body of work. In discussing field and quasi-experiments that did and did not involve randomization, Shadish (2002) observed that while "… nonrandomized experiments continue to be … tarred with a general brush of inferiority" (p. 3), … "across a program of research, all validity types are a high priority, and by the end of a program of research, each should have had its turn in the spotlight" (p. 10).

Additional research questions require attention if the field is to use findings concerning contextual, psychological, and treatment influences on functioning and disability over time. We have discussed how this suggests an urgent need for theoretical development in rehabilitation psychology and the development of valid indicators of participation as a key outcome in rehabilitation and broader efforts to improve the lives of people with disabilities. Wider adoption of a multivariate longitudinal approach will have a salutary effect in this regard. It will help to close the gap between a public health oriented disability model and the narrower clinical rehabilitation model by enhancing the ability to conceptualize treatment as one of several variable classes that influence functioning, disability, and participation over time.

There will be costs to this endeavor (Tucker & Roth, 2006). Investigators may need to upgrade their statistical skills to take advantage of modern data analytic capabilities and to communicate effectively about these approaches. Some approaches will require larger samples, longer studies, more interdisciplinary collaboration, and more multi-center investigations in order to accrue sufficient numbers of participants, given the relatively low base rates of the conditions of concern to rehabilitation psychology, even when functionally defined. We believe the payoffs from these developments will be substantial. When combined with the existing achievements of rehabilitation psychology, they can help position the field to strengthen linkages between science and practice through conceptually guided research on the multi-determined and longitudinal process of adaptation and change among persons living with disabilities.

References

Abrams, DB.; Emmons, KM. Health behavior and health education: The past, present, and future. In: Glanz, K.; Rimer, BK.; Lewis, FM., editors. Health behavior and health education: Theory, research, and practice. San Francisco: Jossey-Bass; 1997. p. 453-478.

Agency for Health Care Policy and Research. Clinical Practice Guideline Development (AHCPR Publication No. 93–0023). Rockville, MD: U.S. Department of Health and Human Services; 1993.

Americans with Disabilities Act. 1990. Pub. L. No. 101–336, 42 U.S.C. § 12101. Retrieved September 11, 2007 from http://www.ada.gov/pubs/ada.htm

Andersen RM. Revisiting the behavioral model and access to medical care: does it matter? Journal of Health and Social Behavior 1995;36:1–10. [PubMed: 7738325]

Atchison TB, Sander AM, Struchen MA, High WM, Roebuck TM, Contant CF, et al. Relationship between neuropsychological test performance and productivity at 1 year following traumatic brain injury. Clinical Neuropsychology 2004;18:249–265.

Benson K, Hartz AJ. A comparison of observational studies and randomized, controlled clinical trials. New England Journal of Medicine 2000;342:1878–1886. [PubMed: 10861324]

Blacker CVR, Mortimore C. Randomized controlled trials and naturalistic data: Time for a change? Human Psychopharmacology 1996;11:353–363.

Bombardier CH, Richards JS, Krause JS, Tulsky D, Tate DG. Symptoms of major depression in people with spinal cord injury: Implications for screening. Archives of Physical Medicine and Rehabilitation 2004;85:1749–1756. [PubMed: 15520969]

Borckardt JJ, Nash MR, Murphy MD, Moore M, Shaw D, O'Neil P. Clinical practice as a natural laboratory for psychotherapy research: A guide to case-based time-series analysis. American Psychologist 2008;63:77–95. [PubMed: 18284277]

Boutron I, Moher D, Altman DG, Schulz KF, Ravaud P. Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: Explanation and elaboration. Annals of Internal Medicine 2008;148:295–309. [PubMed: 18283207]

Brownson, RC.; Baker, EA.; Leet, TL.; Gillespie, KN. Evidence-based public health. New York: Oxford University Press; 2003.

Bruyère S. Policies and practices to promote nondiscrimination across the employment process. American Rehabilitation. in press

Cardol M, De Jong BA, Ward CD. On autonomy and participation in rehabilitation. Disability and Rehabilitation 2002;24:970 –974. [PubMed: 12528676]

Chambless DL, Sanderson WC, Shoham V, Johnson SB, Pope KS, Rits-Christoph P, et al. An update on empirically validated therapies. The Clinical Psychologist 1996;49:5–18.

Chamie M. What does morbidity have to do with disability? Disability and Rehabilitation 1995;17:323–337. [PubMed: 8680032]

Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. New England Journal of Medicine 2000;342:1887–1892. [PubMed: 10861325]

Corrigan, PW.; Mueser, K.; Bond, GR.; Drake, RE.; Solomon, P., editors. Principles and practice of psychiatric rehabilitation: An empirical approach. New York: Guilford; 2007.

Department of Health and Human Services. Healthy people 2010. Washington, DC: Author; 2000.

Dunn DS, Dougherty SB. Prospects for a positive psychology or rehabilitation. Rehabilitation Psychology 2005;50:305–311.

Elkin I, Shea MT, Watkins JT, Imber SD, Sotsky SM, Collins JF, et al. National Institute of Mental Health Treatment of Depression Collaborative Research Program: General effectiveness of treatments. Archives of General Psychiatry 1989;46:971–982. [PubMed: 2684085]

Elliott T. A place for theory in the study of psychological adjustment among persons with neuromuscular disorders: A reply to Livneh and Antonak. Journal of Social Behavior and Personality 1994;9:231–236.

Elliott T, Frank RG. Depression following spinal cord injury. Archives of Physical Medicine and Rehabilitation 1996;77:816–823. [PubMed: 8702378]

Elliott, TR.; Frank, RG. Afterward: Drawing new horizons. In: Frank, RG.; Elliott, T., editors. Handbook of rehabilitation psychology. Washington, DC: American Psychological Association; 2000. p. 645-653.

Elliott, T.; Jackson, WT. Cognitive-behavioral therapy in rehabilitation psychology. In: Freeman, A., editor. Encyclopedia of cognitive behavior therapy. New York: Springer Science and Business Media; 2005. p. 324-327.

Elliott, T.; Kurylo, M.; Rivera, P. Positive growth following an acquired physical disability. In: Snyder, CR.; Lopez, S., editors. Handbook of positive psychology. New York: Oxford University Press; 2002. p. 687-699.

Elliott, T.; Leung, P. Vocational rehabilitation: History and practice. In: Walsh, WB.; Savickas, M., editors. Handbook of vocational psychology. Vol. 3. New York: Lawrence Erlbaum Associates; 2005. p. 319-343.

Elliott, TR.; Warren, AM. Why psychology is important in rehabilitation. In: Kennedy, P., editor. Psychological management of physical disabilities: A practitioner's guide. London: Routledge Press; 2007. p. 16-39.

Fordyce, W. Behavioral methods for chronic pain and illness. St Louis, MO: Mosby; 1976.

Fuhrer MJ. Overview of clinical trials in medical rehabilitation: Impetuses, challenges, and needed future directions. American Journal of Physical Medicine and Rehabilitation 2003;82 (Suppl):S8–S15. [PubMed: 14502033]

Gatchel, RJ. Clinical essentials of pain management. Washington, DC: American Psychological Association; 2005.

Gatchel RJ, Polatin PB, Mayer TG. The dominant role of psychosocial risk factors in the development of chronic low back pain disability. Spine 1995;20:2702–2709. [PubMed: 8747248]

Goodheart, CD.; Kazdin, AE.; Sternberg, RJ., editors. Evidence-based psychotherapy: Where practice and research meet. Washington, D.C.: American Psychological Association; 2006.

Heinemann AW. Putting outcome measurement in context: A rehabilitation psychology perspective. Rehabilitation Psychology 2005;50:6–14.

Hendershot, GE.; Lollar, D. The International Classification of Functioning: Facilitating integrative healthcare. In: Serlin, IA.; DiCowden, MA., editors. Whole Person Healthcare—Volume 1: Humanizing healthcare. Westport, CT: Praeger; 2007. p. 233-248.

Hollon, SD. What qualifies as evidence on which to judge effective practice? Randomized clinical trials. In: Norcross, JC.; Beutler, LE.; Levant, RF., editors. Evidence-based practices in mental health. Washington, DC: American Psychological Association; 2006. p. 96-105.

Hollon SD, DeRubeis RJ, Shelton RC, Amsterdam JD, Salomon RM, O'Reardon JP, et al. Prevention of relapse following cognitive therapy vs. medications in moderate to severe depression. Archives of General Psychiatry 2005;62:417–422. [PubMed: 15809409]

Hollon SD, Jarrett RB, Nierenberg AA, Thase ME, Trivedi M, Rush AJ. Psychotherapy and medication in the treatment of adult and geriatric depression: Which monotherapy or combined treatment? Journal of Clinical Psychiatry 2005;66:455–468. [PubMed: 15816788]

Hollon SD, Thase ME, Markowitz JC. Treatment and prevention of depression. Psychological Science in the Public Interest 2002;3:39–77.

Horn SD. Performance measures and clinical outcomes. Journal of the American Medical Association 2006;296:2731–2732. [PubMed: 17164460]

Horn SD, DeJong G, Ryser DK, Veazie PJ, Teraoka J. Another look at observational studies in rehabilitation research: Going beyond the Holy Grail of the randomized controlled trial. Archives of Physical Medicine and Rehabilitation 2005;86 (Suppl 2):S8–S15. [PubMed: 16373136]

Horn SD, Gassaway J. Practice-based evidence study design for comparative effectiveness research. Medical Care 2007;45 (Suppl 2):S50–S57. [PubMed: 17909384]

Hulley, SB.; Cummings, SR. Designing clinical research. Baltimore: Williams and Wilkins; 1988.

Institute of Medicine. Crossing the quality chasm: A new health system for the 21st century. Washington, DC: National Academies Press; 2001.

Institute of Medicine. Improving the quality of health care for mental health and substance use conditions. Washington, D.C.: National Academies Press; 2005.

Institute of Medicine. The future of disability in America. Washington, DC: National Academies Press; 2007.

Jenicek, M. Foundations of evidence-based medicine. New York: Parthenon Publishing Group; 2003.

Judd CM, Kenny DA, McCelland GH. Estimating and testing mediation and moderation in within-subjects designs. Psychological Methods 2001;6:115–134. [PubMed: 11411437]

Kazdin, AE. Single-case research designs: Methods for clinical and applied settings. New York: Oxford University Press; 1982.

Kazdin AE. Mediators and mechanisms of change in psychotherapy research. Annual Review of Clinical Psychology 2007;3:1–27.

Kraemer HC, Stice E, Kazdin A, Offord D, Kupper D. How do risk factors work together: Mediators, moderators, and independent, overlapping, and proxy risk factors. American Journal of Psychiatry 2001;158:848–856. [PubMed: 11384888]

Kuhn, TS. The structure of scientific revolutions. Vol. 2. Chicago: University of Chicago Press; 1970.

Langhorne P, Duncan P. Does the organization of postacute stroke care really matter? Stroke 2001;32:268–274. [PubMed: 11136947]

Leichsenring F. Randomized controlled versus naturalistic studies: A new research agenda. Bulletin of the Menninger Clinic 2004;68:137–151. [PubMed: 15262616]

Linden W, Stossel C, Maurice J. Psychosocial interventions for patients with coronary artery disease: A meta-analysis. Archives of Internal Medicine 1996;156:745–752. [PubMed: 8615707]

Lollar DJ. Secondary conditions among adolescents with spina bifida. European Journal of Pediatric Surgery 1997;7 (Suppl 1):55.

Lollar, D. Clinical dimensions of secondary conditions. In: Simeonsson, RJ.; McDivitt, LN., editors. Issues in disability and health: The roles of secondary conditions and quality of life. Chapel Hill: Univ. of North Carolina, FPG Child Development Center; 1999.

Lollar DJ, Crews JE. Redefining the role of public health in disability. Annual Review of Public Health 2003;24:195–208.

Lollar D, Simeonsson R. Diagnosis to function: Classification for children and youths. Journal of Developmental Behavior and Pediatrics 2005;26:323–330.

MacKenzie, BB. Behaviourism and the limits of scientific method. Atlantic Highland, NJ: Humanities Press; 1977.

Marson DC, Dreer LE, Krzywanski S, Huthwaite JS, Devivo MJ, Novack TA. Impairment and partial recovery of medical decision-making capacity in traumatic brain injury: A 6-month longitudinal study. Archives of Physical Medicine and Rehabilitation 2005;86:889–895. [PubMed: 15895333]

McMurray JJ, O’Meara E. Treatment of heart failure with spironolactone: Trial and tribulations. New England Journal of Medicine 2004;351:526–528. [PubMed: 15295043]

Meadows, M. FDA Consumer. 2002. The FDA’s drug review process: Ensuring drugs are safe and effective. (FDA Publication no. 02–3242)

Mechanic, D. Medical sociology. Vol. 2. New York: Free Press; 1978.

Mendes de Leon CF, Czajkowski SM, Freedland KE, Bang H, Powell LH, Wu C, et al. The effect of a psychosocial intervention and quality of life after acute myocardial infarction: The Enhancing Recovery in Coronary Heart Disease (ENRICHD) clinical trial. Journal of Cardiopulmonary Rehabilitation and Prevention 2006;26:9–13.

Miller FG, Brody H. Clinical equipoise and the incoherence of research ethics. Journal of Medicine and Philosophy 2007;32:151–165. [PubMed: 17454420]

Miller, WR.; Rollnick, S. Motivational interviewing: Preparing people for change. Vol. 2. New York: Guilford; 2002.

Moher D, Schulz KF, Altman D. The CONSORT statement: Revised recommendations for improving the quality of reports of parallel-group randomized trials. Journal of the American Medical Association 2001;285:1987–1991. [PubMed: 11308435]

Morrisey, M. Price sensitivity in health care: Implications for health care policy. Washington, D.C.: National Federation of Independent Business Foundation; 1992.

National Center for Health Statistics. Health, United States, 2006, with chartbook on trends in the health of Americans. Hyattsville, MD: Author; 2006.

Niemeier JP, Kreutzer JS, Taylor LA. Acute cognitive and neurobehavioural intervention for individuals with acquired brain injury: preliminary outcome data. Neuropsychological Rehabilitation 2005;15:129–146. [PubMed: 16353506]

Norcross, JC. Psychotherapy relationships that work: Therapists, contributions and responsiveness to patients. New York: Oxford University Press; 2002.

Norcross, JC.; Beutler, LE.; Levant, RF., editors. Evidence-based practices in mental health: Debate and dialogue on fundamental questions. Washington, D.C.: American Psychological Association; 2006.

Olkin R. Making research accessible to participants with disabilities. Journal of Multicultural Counseling & Development 2002;8:130–137.

Olkin R, Pledger C. Can disability studies and psychology join hands? American Psychologist 2003;58:296–304. [PubMed: 12866396]

Omenn GS, Goodman GE, Thomquist MD, Balmes J, Cullen MR, Glass A, et al. Effects of a combination of beta-carotene and vitamin A on lung cancer and cardiovascular disease. New England Journal of Medicine 1996;334:1150–1155. [PubMed: 8602180]

Pepper, SC. World hypotheses. Berkeley, CA: University of California Press; 1972.

Pitt B, Zannad F, Remme WJ, Cody R, Castaigne A, Perez A, Palensky J, Wittes J. The effect of spironolactone on morbidity and mortality in patients with severe heart failure. New England Journal of Medicine 1999;341:709–17. [PubMed: 10471456]

Pope, AM.; Tarlov, AR., editors. Disability in America: Toward a national agenda for prevention. Washington, DC: National Academies Press; 1991.

Public Health Information & Data Tutorial. Hierarchy of evidence. 2006. Retrieved February 15, 2008 from http://phpartners.org/tutorial/04-ebph/2-keyConcepts/4.2.7.html

Purtilo, RB.; Haddad, AM. Health professional and patient interaction. Vol. 7. Philadelphia, PA: W.B. Saunders; 2007.

Ravesloot C, Seekins T, Walsh J. A structural analysis of secondary conditions of primary physical disabilities. Rehabilitation Psychology 1997;42:3–16.

Reed, GM. What qualifies as evidence of effective practice: Clinical expertise. In: Norcross, JC.; Beutler, L.; Levant, RF., editors. Evidence-based practices in mental health: Debate and dialogue on the fundamental questions. Washington, DC: American Psychological Association; 2006. p. 13-23.

Reed, GM.; Eisman, E. Uses and misuses of evidence: Managed care, treatment guidelines, and outcomes measurement in professional practice. In: Goodheart, CD.; Kazdin, AE.; Sternberg, RJ., editors. Evidence-based psychotherapy: Where research and meet. Washington, DC: American Psychological Association; 2006. p. 13-25.

Reed GM, Lux JB, Bufka LF, Trask C, Peterson DB, Stark S, Threats TT, Jacobson JW, Hawley JA. Operationalizing the International Classification of Functioning, Disability and Health (ICF) in clinical settings. Rehabilitation Psychology 2005;50:122–131.

Rosen L, Manor O, Engelahrd E, Zucker D. In defense of the randomized controlled trial for health promotion research. American Journal of Public Health 2006;96:1181–1186. [PubMed: 16735622]

Rosenberg M, Plakeney P, Robert R, Thomas C, Holzer C, Meyer W. Quality of life of young adults who survived pediatric burns. Journal of Burn Care and Rehabilitation 2006;27:773–778.

Rosenberg L, Robert R, Thomas C, Holzer C, Blakeney P, Meyer W. Assessing potential suicide risk of young adults burned as children. Journal of Burn Care and Rehabilitation 2006;27:779–785.

Sackett, DL.; Straus, SE.; Richardson, WS.; Rosenberg, W.; Haynes, RB. Evidence based medicine: How to practice and teach EBM. Vol. 2. London: Churchill Livingstone; 2000.

Sacks H, Chalmers I, Smith H Jr. Randomized versus historical controls for clinical trials. American Journal of Medicine 1982;72:2330–240.

Scherer, MJ.; Coombs, FK.; Merbitz, NH.; Merbitz, CT. Policy issues in evaluating and selecting assistive technology and other resources for persons with disabilities. In: Hagglund, KJ.; Heinemann, AW., editors. Handbook of Applied Disability and Rehabilitation Research. New York: Springer; 2006. p. 215-242.

Seligman MEP. The effectiveness of psychotherapy: The *Consumer Reports*study. American Psychologist 1995;50:965–974. [PubMed: 8561380]

Shadish WR. Revisiting field experimentation: Field notes for the future. Psychological Methods 2002;7:3–18. [PubMed: 11928889]

Shadish WR, Navarro AM, Matt GE, Phillips G. The effects of psychological therapies under clinically representative conditions: A meta-analysis. Psychological Bulletin 2000;126:512–529. [PubMed: 10900994]

Shrout PE, Bolger N. Mediation in experimental and nonexperimental studies: New procedures and recommendations. Psychological Methods 2002;4:422–445. [PubMed: 12530702]

Singer, JD.; Willet, JB. Applied longitudinal data analysis: Modeling change and event occurrence. New York: Oxford University Press; 2003.

Spring B, Pagoto S, Kaufmann PG, Whitlock EP, Glasgow RE, Smith TW, Trudeau KH, Davison KW. Invitation to a dialogue between researchers and clinicians about evidence-based behavioral medicine. Annals of Behavioral Medicine 2005;30:125–137. [PubMed: 16173909]

Tanenbaum SJ. Evidence-based practice as mental health polity: Three controversies and a caveat. Health Affairs 2005;24:163–173. [PubMed: 15647227]

Tate DG, Findley T Jr, Dijkers M, Nobunaga AI, Karunas RB. Randomized clinical trials in medical rehabilitation research. American Journal of Physical Medicine and Rehabilitation 1999;78:486–499. [PubMed: 10493460]

Tate DG, Forchheimer M, Krause JS, Meade M, Bombardier CH. Patterns of alcohol and substance use and abuse in persons with spinal cord injury: Risk factors and correlates. Archives of Physical Medicine and Rehabilitation 2004;85:1837–1847. [PubMed: 15520979]

Tate DG, Kalpakjian CZ, Forchheimer M. Quality of life issues in individuals with SCI. Archives of Physical Medicine and Rehabilitation 2002;83 (Suppl 2):S18–S25. [PubMed: 12474168]

Taub, E. Somatosensory deafferentation research with monkeys: Implications for rehabilitation medicine. In: Ince, LP., editor. Behavioral psychology in rehabilitation medicine: Clinical applications. New York: Williams & Wilkins; 1980. p. 371-401.

Taub E, Miller NE, Novack TA, Cook EW, Fleming WC, Nepomuceno CS, et al. Technique to improve chronic motor deficit after stroke. Archives of Physical Medicine and Rehabilitation 1993;74:347–354. [PubMed: 8466415]

Taub E, Uswatte G. Constraint-Induced Movement therapy: Answers and questions after two decades of research. NeuroRehabilitation 2006;21:93–95. [PubMed: 16917156]

Taub E, Uswatte G, Elbert T. New treatments in neurorehabilitation founded on basic research. Nature Reviews Neuroscience 2002;83:228 –236.

Taub E, Uswatte G, King DK, Morris D, Crago J, Chatterjee A. A placebo controlled trial of Constraint-Induced Movement therapy for upper extremity after stroke. Stroke 2006;37:1045–1049. [PubMed: 16514097]

Taub E, Uswatte G, Mark VW, Morris DM. The learned nonuse phenomenon: Implications for rehabilitation. Europa Medicophysica 2006;42:241–255. [PubMed: 17039223]

Taylor SE. Adjustment to threatening events: A theory of cognitive adaptation. American Psychologist 1983;38:1161–1173.

Taylor SE, Brown JD. Illusion and well-being: A social psychological perspective on mental health. Psychological Bulletin 1988;110:193–210. [PubMed: 3283814]

Taylor SE, Kemeny ME, Aspinwalk LG, Schneider SC, Rodriguez R, Herbert M. Optimism, coping, psychological distress, and high-risk sexual behavior among men at risk for AIDS. Journal of Personality and Social Psychology 1992;63:460–473. [PubMed: 1403625]

Taylor SE, Kemeny ME, Reed GM, Bower JE, Gruenewald TL. Psychological resources, positive illusions, and health. American Psychologist 2000;55:99–109. [PubMed: 11392870]

Tucker, JA. Changing addictive behavior: Historical and contemporary perspectives. In: Tucker, JA.; Donovan, DM.; Marlatt, GA., editors. Changing addictive behavior: Bridging clinical and public health perspectives. New York: Guilford; 1999. p. 3-44.

Tucker, JA.; King, MP. Resolving alcohol and drug problems: Influences on addictive behavior change and help-seeking processes. In: Tucker, JA.; Donovan, DM.; Marlatt, GA., editors. Changing addictive behavior: Bridging clinical and public health perspectives. New York: Guilford; 1999. p. 97-126.

Tucker JA, Roth DL. Extending the evidence hierarchy to enhance evidence-based practice for substance use disorders. Addiction 2006;101:918–932. [PubMed: 16771885]

Van Brakel WH, Anderson AM, Mutatkar RK, Bakirtzief Z, Nicholls PG, Raju MS, Das-Pattanayak RK. The Participation Scale: measuring a key concept in public health. Disability and Rehabilitation 2006;28:193–203. [PubMed: 16467054]

Victoria CG, Habicht JP, Bryce J. Evidence-based public health: Moving beyond randomized trials. American Journal of Public Health 2004;94:400–405. [PubMed: 14998803]

Walker WC, Marwitz JH, Kreutzer JS, Hart T, Novack TA. Occupational categories and return to work after traumatic brain injury: A multicenter study. Archives of Physical Medicine and Rehabilitation 2006;87:1576–1582. [PubMed: 17141636]

Wassertheil-Smoller S, Hendrix SL, Limacher M, Heiss G, Kooperberg C, Baird A, et al. Effect of estrogen plus progestion on stroke in postmenopausal women: The Women's Health Initiative: A randomized trial. Journal of the American Medical Association 2003;289:2673–2684. [PubMed: 12771114]

Weinstein ND. Why it won't happen to me: Perceptions of risk factors and susceptibility. Health Psychology 1984;3:431–457. [PubMed: 6536498]

Westen D, Novotny CM, Thompson-Brenner H. Empirical status of empirically supported psychotherapies: Assumptions, findings, and reporting in controlled clinical trials. Psychological Bulletin 2004;130:631–663. [PubMed: 15250817]

Whiteneck G, Forchheimer MB, Krause JS. Quality of life and health in the last years after spinal cord injury. Topics in Spinal Cord Injury Rehabilitation 2007;12:77–90.

Whiteneck G, Meade MA, Dijkers M, Tate DG, Bushnik T, Forchheimer MB. Environmental factors and their role in participation and life satisfaction after spinal cord injury. Archives of Physical Medicine and Rehabilitation 2004;85:1793–1803. [PubMed: 15520974]

Wolf SL, Winstein CJ, Miller JP, Taub E, Uswatte G, Morris D, et al. Effect of Contraint-Induced Movement Therapy on upper extremity function 3 to 9 months after stroke: The EXCITE randomized clinical trial. Journal of the American Medical Association 2006;296:2095–2104. [PubMed: 17077374]

World Health Organization. Health promotion evaluation recommendations to policy makers. Report of the WHO European Working Group on health promotion evaluation. Geneva: Author; 1998.

Word Health Organization. International Classification of Functioning, Disability, and Health. Geneva: Author; 2001.

Wright, GN. Total rehabilitation. Boston: Little, Brown; 1980.

Table 1

Hierarchical classification of evidence informing treatment efficacy

I.	Randomized controlled clinical trials.
II.	Well designed trials without randomization. Non-randomized trials or those with high alpha and beta errors.
III.	Analytical observational studies.
IV.	Multiple time series or place comparisons, uncontrolled (‘natural’) experiments.
V.	Expert opinions, descriptive occurrence studies, case reports, case series reports.

Note. From Jenicek (2003, p. 34).

Table 2

Efficacy vs. effectiveness studies of treatment effects

	Efficacy studies	Effectiveness studies
Purpose	detect the “true effects” of treatment by reducing extraneous variance	assess treatment effectiveness in usual care settings
Typical characteristics	• homogeneous samples • random assignment, blinded • manualization, standardized treatment, fidelity checks	• heterogeneous samples • circumstances leading to and surrounding care studied • compliance measured, not forced
Better suited for studying	• technical interventions with acute effects (e.g., surgery, pharmacotherapy) • interventions with effects that are fairly independent of context • interventions amenable to fine adjustments to improve outcomes	• psycho-social problems and interventions distributed through time (e.g., functional recovery after injury) • problems and interventions that are influenced by context • care studied as a process, not just as an outcome
Validity and Inference	Higher internal validity but lower generalization potential and capacity to inform science-to-practice linkages	Lower internal validity but higher generalization potential and capacity to inform science-to-practice linkages

Note. Adapted from Tucker (1999, p. 25).