



The Challenge of Evidence in Disability and Rehabilitation Research and Practice

A Position Paper

Mark V. Johnston, PhD
Gregg C. Vanderheiden, PhD
Marianne D. Farkas, ScD
E. Sally Rogers, ScD
Jean Ann Summers, PhD
John D. Westbrook, PhD

For the Task Force on Standards of Evidence and Methods

NATIONAL CENTER FOR THE
DISSEMINATION OF DISABILITY RESEARCH

Advancing Research, Improving Education



The Challenge of Evidence in Disability and Rehabilitation Research and Practice: A Position Paper

Introduction to the NCDDR Task Force Papers

The National Center for the Dissemination of Disability Research (NCDDR) has established three task forces to assist the project in analyzing, understanding, and commenting on features of the evidence production process within the disability and rehabilitation research context. See www.ncddr.org/kt_task_forces.html

- Task Force on Standards of Evidence and Methods
- Task Force on Systematic Review and Guidelines
- Task Force on Knowledge Translation/Knowledge Value Mapping

Each task force is comprised of senior researchers with current or recent experience in conducting NIDRR-sponsored research activities. Each task force is charged with developing positions and statements that are relevant in light of current circumstances.

This paper was developed as a group effort of the National Center for the Dissemination of Disability Research (NCDDR) Task Force on Standards of Evidence and Methods (TFSE). Concerns and suggestions of all members have been taken into consideration in the development of this paper. All issues, however, could not be fully addressed in this initial paper. Subsequent papers will address key issues related to this topic area. Task Force members contributing to this paper include:

Mark V. Johnston, PhD; Gregg C. Vanderheiden, PhD; Marianne D. Farkas, ScD; E. Sally Rogers, ScD; Jean Ann Summers, PhD; and John D. Westbrook, PhD

TASK FORCE FACILITATOR:

Mark V. Johnston, PhD

AFFILIATION:

University of Wisconsin-Milwaukee, WI

CORRESPONDENCE:

National Center for the Dissemination of Disability Research (NCDDR)

SEDL

4700 Mueller Blvd.

Austin, TX 78723-3081

ncddr@sedl.org

DISCLOSURE:

The Task Force on Standards of Evidence and Methods is sponsored by the National Center for the Dissemination of Disability Research (NCDDR) and funded by the National Institute on Disability and Rehabilitation Research (NIDRR).

RECOMMENDED CITATION:

Johnston, M. V., Vanderheiden, G. C., Farkas, M. D., Rogers, E. S., Summers, J. A., & Westbrook, J. D., for the NCDDR Task Force on Standards of Evidence and Methods. (2009). *The challenge of evidence in disability and rehabilitation research and practice: A position paper*. Austin, TX: SEDL.

View the paper online at www.ncddr.org/kt/products/ftpapers/tfse_challenge/

The Challenge of Evidence in Disability and Rehabilitation Research and Practice: A Position Paper

Can evidence beneficially shape information resources and services for people with disabilities? What constitutes “best evidence” on interventions for people with disabilities? Do we need to develop specific evidence standards to identify best evidence on interventions for people with disabilities?

This paper states the position of the NCDDR Task Force on Standards of Evidence and Methods (TFSE) regarding the need for (a) the thoughtful determination of research evidence on the basis of both the rigor of the research and the relevance of the research to the lives of people with disabilities; and (b) systems that facilitate our ability, on a timely basis, to describe what the best available evidence is in response to specific topical questions in disability and rehabilitation.

The primary focus of this paper is on evidence for interventions in the field of disability and rehabilitation (D&R). Evidence issues related to D&R interventions concern all people with disabilities and involve both research and development, as both are extremely important. The specific objectives of this paper are the following:

- To clarify what is meant by the term *evidence* and to describe the nature of the contemporary systems used to identify and evaluate evidence in intervention research
- To identify the challenges in meeting contemporary standards of evidence in the field of D&R interventions
- To propose next steps for examining related issues and for taking action to promote the availability of evidence-based services and information in the field of D&R interventions

The Challenge of Evidence

Evidence and Contemporary Systems for Grading Evidence

Few researchers would disagree with the proposition that D&R policies and practices should be grounded in evidence. The issue is how that evidence should be identified, evaluated, and synthesized. What standards and methods should be applied to evaluate the strength of scientific evidence used to inform practices and policies for people with disabilities?

Evidence, for the purpose of this paper, refers to the knowledge that connects research to practice. Over the years, an increasing emphasis on evidence has led to a movement for evidence-based practice (EBP). Emerging first in the health care industry, EBP has since swept into a number of other professional fields, including D&R.

In the field of D&R, EBP involves using the *best available evidence*—integrated with clinical expertise and the values and experiences of people with disabilities and other stakeholders—to guide decisions about clinical and community practices. In this paper, we define D&R *practices*, or *interventions*, as systematic actions, programs, treatments, devices, or environmental changes designed to benefit, either directly or indirectly, individuals or groups with disabilities. In a clinic or home, D&R interventions usually focus on individuals, although small groups (e.g., family) may be treated as well. D&R interventions may also target larger units, such as classrooms, companies, or communities

(e.g., to increase physical accessibility, to alter attitudes, to effect universal design, or to improve policies). Primary domains of concern for D&R interventions include participation and community living, employment, health and function, and technology for access and function (Federal Register, 2006).

To support EBP, many professional organizations have developed detailed evidence grading systems for use mainly in evaluating and synthesizing intervention studies (e.g., Edlund, Gronseth, So, & Franklin, 2004; Guyatt et al., 2008; Higgins & Green, 2006; Institute of Medicine, 2008; Sackett, Straus, Richardson, Rosenberg, & Haynes, 2000). While a variety of rankings are used, evidence is commonly graded on a scale from Level 1, the strongest evidence, to Levels 4 or 5, the weakest evidence.

Such evidence grading systems are increasingly being used in the field of D&R to evaluate studies of clinical and community practices. For example, several grading systems are available for use in selecting the best evidence to answer clinical questions in systematic reviews and meta-analyses. These systems include the Cochrane Collaboration, Campbell Collaboration, Agency for Healthcare Research and Quality (AHRQ), and Grading of Recommendations Assessment, Development and Evaluation (GRADE), as well as grading systems from medical societies such as the American Academy of Neurology, also recommended by the American Congress of Rehabilitation Medicine

The Challenge of Evidence

(Atkins et al., 2004; Edlund et al., 2004; Guyatt et al., 2008; Higgins & Green, 2006; Institute of Medicine, 2008; West et al., 2002).

Evidence grading systems assess and rank the quality of research studies on the basis of pre-established criteria, or standards of evidence, which go beyond dichotomies such as good or bad and rigorous or nonrigorous. Virtually all evidence grading systems for studies of intervention efficacy address the following aspects of research quality (West et al., 2002):

- Randomization, and in some cases other methods of evaluating the comparability of control groups
- Blinding, and in some cases other methods of avoiding measurement biases, attrition, or losses
- Statistical conclusion validity, including the size of statistical confidence intervals

For example, an overarching purpose of using evidence grading systems is to avoid or minimize bias, including not only technical biases in research procedures but also biases associated with self-interest, financial interest, or social pressure to express certain opinions regardless of the scientific data.

In addition to issues of research quality, evidence grading systems (e.g., GRADE) in the field of D&R increasingly consider the strength of recommendations and the relevance of evidence to individuals' needs and values. Such issues of practical application assess the external validity of the evidence, and are just as important for

evidence standards and methods to address as issues of research quality, which assess the internal validity of the evidence. For this reason, future Task Force papers will examine further both the research quality and application sides of the evidence bridge.

The Challenge of Evidence in Disability and Rehabilitation

EBP is quickly becoming the preferred approach for guiding D&R professionals in rendering services to individuals with disabilities. However, contemporary evidence standards and methods pose a number of challenges for the field of D&R. For one, the evidence standards and methods used in many systematic reviews and meta-analyses identify few Level 1 studies of D&R interventions or programs (Johnston, Sherer, & Whyte, 2006). Many systematic reviews and meta-analyses include only randomized controlled trials (RCTs), widely recognized as the most rigorous method of testing intervention efficacy. For this reason, some recently published reviews have reported finding very little or no evidence. Although such results may reflect a scarcity of well-controlled D&R trials rather than a lack of intervention effectiveness, findings of an absence of evidence pose a serious and ongoing challenge to the field of D&R.

The shortage of Level 1 clinical trials in D&R is due in large part to the nature and scope of the field. In both research and practice, D&R is an exceptionally wide, multidisciplinary field involving biological, psychological,

social, economic, legal, and environmental factors related to disability. The field's mission entails the commitment to help people with disabilities "perform activities of their choice" and "to expand society's capacity to provide full opportunities and accommodations for its citizens with disabilities" (Federal Register, 2006). This vast scope of concern includes social integration, employment, independent living, health, and enabling technology. Although basic scientific standards and methods can be applied to D&R, multiple standards and methods are needed to discern the best evidence for the wide and heterogeneous problems and interventions addressed in D&R research and practice.

The nature of D&R presents significant challenges to knowledge development and evidence identification, including the following:

- *Great breadth and complexity.* Conceptually, disability involves the interaction of a person with a wide range of complex factors in the environment (World Health Organization, 2001). In both research and practice, some D&R interventions target health or biological functions; others target skills, feelings, or behaviors; and still others target aspects of the social or physical environment that limit people with disabilities (e.g., attitudes of employers or physical accessibility).
- *Emphasis on empowering people with disabilities.* D&R research involves a commitment to a participatory approach that includes people with disabilities as decision makers throughout the process.

This approach requires research designs and methodologies that appropriately and effectively allow for such participation. Although critical to ensuring that the research is relevant to the lives and values of people with disabilities, these designs and methodologies may be considered less rigorous under most current evidence grading methods.

- *Small sample sizes.* Although disability is common, affecting the majority of people at some point, it is also extremely diverse. Interventions typically must be highly individualized, or client centered, and tailored to particular configurations of impairment or to personal and contextual factors. This diversity and need for customization often result in small samples for studies at any one local site.
- *Difficulty or impossibility of complete blinding and placebo control.* For many personalized therapies, the client and therapist need to be aware of the intervention involved. For example, researchers cannot hide from clients the presence of an assistive device or a guide dog.
- *Difficulty in defining an ethical and practical control group.* RCTs are comparatively new to D&R and a departure from its research tradition. Practitioners and people with disabilities are apprehensive about the denial of services for control groups in RCTs.
- *Need for enabling technology, including assistive devices and environmental*

modifications, to improve a disabled person's chosen activities or quality of life. Existing evidence grading systems do not address all of the research methods used to evaluate assistive technology or universal design for accessibility and successful use.

- *Funding levels that are adequate for pilot studies, intervention development, or early stage clinical trials but not for truly rigorous effectiveness studies.* D&R is widely perceived as involving issues related to clinical service delivery and advocacy (Field, Jette, & Martin, 2007) rather than issues related to research and development. As a result, funding levels for research are often inadequate for rigorous scientific inquiries using a large, multisite RCT design.
- *Need to address issues within large social systems that involve consideration for the social, physical, and/or economic environment.* Many of the major issues in D&R concern large social systems that cannot be manipulated experimentally (e.g., universal design, accessibility, public attitudes, legal rights, effects of culture, economic factors). These contextual effects are not readily incorporated into current evidence grading systems.

Current EBP standards and methods for D&R were derived from evidence-based medicine (EBM) and optimized for well-funded studies of well-defined, individual-level clinical interventions, such as pharmaceutical agents, which should be tested using blinded RCTs.

For many of the current research problems in D&R, however, the usual or optimal solution will not be a large RCT. The best research design is not always the largest or most rigorous one possible; it is rather the one that will most advance knowledge on the basis of the state of prior research and development and resource constraints.

To identify the best evidence for many D&R practices, EBP standards and methods for systematic reviews need to be sensitive to non-RCT evidence and to recognize classes of interventions for which RCTs, though occasionally possible and worthwhile, are not expected to be the usual or standard source of evidence (e.g., most assistive or enabling technology). For example, methods of controlling for differences between experimental and control groups other than RCTs exist, and D&R evidence grading systems should incorporate those methods (Institute of Medicine, 2008; Johnston, Ottenbacher, & Reichardt, 1995; Schneider, Carnoy, Kilpatrick, Schmidt, & Shavelson, 2007; Victora, Habicht, & Bryce, 2004; West et al., 2008). At the same time, evidence standards and methods in D&R should continue to support the need to develop and test new interventions using the most rigorous methods, including RCTs, whenever appropriate (Johnston & Case-Smith, 2009).

D&R can also benefit from the study of best practices in other fields facing similar challenges with EBP and then adopt or adapt those practices that apply. For instance, researchers

can overcome problems of small sample size by using multicenter, collaborative networks or simply avoid the problem by focusing on issues common to a large number of people. Complex, vague intervention strategies can be outlined, and procedures developed to ensure intervention fidelity. In public health, nursing, psychology, and other fields, hundreds of RCTs have been mounted to study multifaceted community and behavioral interventions (e.g., chronic disease self-management). In other studies, mixed methods have been applied to understand problems of qualitative complexity and context. Ethical clinical trials are mounted in many fields (e.g., by comparing a promising new-or-improved but unproven intervention to treatment as usual), and sophisticated correlational and quasi-experimental research designs can be used when randomized control is not feasible (Johnston et al., 1995; Schneider et al., 2007; West et al., 2008). Fields as diverse as psychology and public health have employed participatory research strategies to enhance study relevance and success (Viswanathan et al., 2004).

Next Steps

To advance EBP in D&R, we need to redefine the field's evidence standards and methods for intervention-based research. Guidelines for evidence development and application should address scientific research quality, relevance to the needs and values of people with disabilities, and applicability to practice. Many evidence grading systems currently available focus primarily on research quality (internal validity),

creating a need to expand the grading systems' relevance to the needs of people with disabilities and the practical applicability of evidence (external validity). One of the next steps related to relevance should involve refining how D&R professionals measure the true needs, views, and desires of people with disabilities.

In addition, because of the breadth of D&R interventions, the field needs to consider developing several evidence grading systems pertaining to specific intervention focuses. Examples of those focuses include interventions involving assistive or enabling technology and devices; behavioral and activity-based interventions; and interventions addressing environmental factors such as social, attitudinal, and physical environments in a larger community or societal context. Similarly, a need exists for developing evidence grading systems for nonintervention research and development in D&R, such as measurement, prognosis, and technology development processes.

The challenges identified in this paper need to be addressed in further detail, and possible solutions proposed based on reviews of current best knowledge. This process should involve consensus development among research and evidence experts both from within and outside of D&R, including representatives of people with disabilities and other stakeholders. D&R research encompasses widely varying professional traditions, and considerable work is required to reach consensus on quality indicators and useful educational materials for the various problems

