



Published in final edited form as:

Arch Phys Med Rehabil. 2012 August ; 93(8 0): S101–S110. doi:10.1016/j.apmr.2011.11.040.

Advancing the Evidence Base of Rehabilitation Treatments: A Developmental Approach

John Whyte, MD, PhD and A.M. Barrett, MD

Moss Rehabilitation Research Institute, Elkins Park, PA (Whyte); Kessler Foundation Research Center, West Orange, NJ (Barrett); and the University of Medicine and Dentistry, New Jersey—NJ Medical School, Newark, NJ (Barrett)

Abstract

Translational research refers to the development of new scientific discoveries into evidence-based treatments for human diseases and conditions. This developmental process requires that a number of scientific, as well as social and psychological obstacles, be overcome during a sequence of research stages that address different goals. Rehabilitation, like other biomedical disciplines, requires this kind of developmental process. For a variety of reasons, however, development of rehabilitation treatments is less linear than the familiar phases of pharmaceutical research. In addition, research on treatments intended to address impairments (body structure/function, in terms of the *International Classification of Functioning, Disability and Health*), faces the challenge of determining the likely impact of an impairment-level treatment on the multifaceted activities and aspects of participation that are the typical goals of rehabilitation treatments. This article describes the application of treatment theory and enablement theory to the development of new impairment-based treatments, and examines similarities and differences between the developmental sequence needed for rehabilitation treatment research versus pharmaceutical research in other areas of medicine.

Keywords

Rehabilitation; Translational medical research

Translational research refers to the developmental process by which basic science discoveries are developed into effective treatments which affect the public health. This flow from the bench to the bedside has been criticized for the slow pace by which our investment in emerging scientific advances actually impacts clinical treatment.¹ In curative biomedical research, knowledge translation faces 2 obstacles which greatly slow the process²: determining how to adapt treatment mechanisms identified in tissue or animal models to human application (fig 1, #1); and achieving adoption of new proven-effective treatments by practicing clinicians (see fig 1, #2).³ Rehabilitation treatment research, which aims primarily to enhance human activities and participation rather than eradicate disease, brings us to consider a third translational obstacle (see fig 1, #3)—the challenge of determining the degree to which a patient's improvement in a treated impairment is likely to result in meaningful improvements in activities and participation.

© 2012 by the American Congress of Rehabilitation Medicine

Reprint requests to John Whyte, MD, PhD, Moss Rehabilitation Research Institute, 50 Township Line Rd, Elkins Park, PA 19027, jwhyte@einstein.edu.

No commercial party having a direct financial interest in the results of the research supporting this article has or will confer a benefit on the authors or on any organization with which the authors are associated.

In this article, we will discuss the unique challenges in conducting translational research in rehabilitation. In doing so, we will address several themes. First, we will emphasize the importance of a phased developmental process for transforming emerging scientific discoveries into effective rehabilitation interventions.⁴⁻⁷ Second, we will address the relevance of 2 distinct bodies of theory to this developmental process—treatment theory (relevant to the direct effects of treatments) and enablement theory (relevant to the impact of treatment-induced changes on more distal aspects of function).⁵ Lastly, we will discuss similarities and differences between the phases of translational research typically seen in drug development, and those that are needed to arrive at effective rehabilitation treatments. Throughout our discussion in this article, we will be concerned with the construct and internal validity of treatment research especially with respect to the treatment's effects on meaningful functional outcomes: how to leverage study design, analysis, and interpretation of research results so as to maximize our rehabilitation research investment, especially relevant when federal research budgets are limited.

It is worth noting that treatments investigated in rehabilitation research might target any of the *International Classification of Functioning, Disability and Health* (ICF) levels (body structure, body function, activity, participation).⁸ However, in this article we focus primarily on treatments that target body structure/function (ie, impairments), because the process of translating basic discovery at the molecular, cellular, and systems level into rehabilitation treatments at the level of the person with limited function and independence is more likely to focus on these levels. Note that we include rehabilitation treatments as not only experience-based treatments delivered by rehabilitation therapists (eg, gait training), but also medications, surgical interventions, equipment, and devices that are applied with the purpose of enhancing function. Treatments directly targeting activities and participation (eg, task and environmental modifications, policy changes) may be more likely to draw on scientific domains such as the psychology of learning or sociology of disability. Nevertheless, many of the research phases discussed here apply to any treatment under development.

CLINICAL RESEARCH PROCEEDS IN PHASES

There are several distinct tasks involved in translating scientific knowledge into effective treatments in use by practicing clinicians. These tasks can be grouped into research development phases, with each phase having a distinct goal—these goals, in turn, are typically addressed by different research methods. The relatively standard developmental trajectory of pharmaceutical research illustrates this fact (table 1). In exploratory or phase 0 research stages, drug discovery births the treatment concept to be explored in subsequent research.⁹ The drug of interest may arise from targeted drug design and preclinical tissue and animal studies; from observation of the effects of naturally occurring chemicals; from observation of fortuitous effects of already marketed drugs; or from unexpected results of research on other, more developed treatments. In phase I, a drug of interest is given to humans to assess specific safety and pharmacokinetic properties. This phase, often conducted with small numbers of healthy subjects, typically involves administering different doses of the drug to establish dose-limiting side effects and optimal dose and schedule. Safe doses of the drug are then given to human subjects with a clinical condition in phase II. This phase, often conducted in a single center, attempts to establish converging support for the drug mechanism suggested at the idea inception stage, and whether it appears feasible to use and worthy of further study. Phase III is devoted to large scale controlled efficacy studies, typically involving multiple clinical testing sites, and ideally results in approval of the drug for marketing. Although phase III involves large numbers of patients and geographic heterogeneity, the patients tested and the clinical researchers who treat them are still highly selected. Thus, phase IV, a less formally organized effectiveness research phase, is devoted

to assessing the impact of the drug as actually used in practice by more varied clinicians on more varied patients, and establishing differences in benefits and adverse effects among different patient subgroups, such as those with more comorbidities and complicating factors.^{10,11}

Although each research phase focuses on a different purpose, the rigor of the design at each phase can also vary which, in turn, determines the strength of the evidence that supports progression to the next phase. Regulatory agencies, such as the Food and Drug Administration, operationalize the approval for movement from phase to phase by requiring certain study designs (eg, placebo controlled trial), and evidence review criteria, such as those used by the American Association of Neurology,^{12,13} seek to formalize evidence quality criteria particularly for moving treatments into large-scale clinical adoption. However, the research design that is sufficiently rigorous to advance knowledge to the next phase is not mapped cleanly to the research phase itself, particularly in rehabilitation research, where testing of nondrug treatments is common.

Phasing in pharmaceutical research is intended to minimize risks to subjects, and ensure that the enormous resources required for multicenter controlled trials are reserved for the most promising treatments. It also, one hopes, develops a feasible treatment regimen, establishes the frequency of adverse treatment effects, and identifies those subjects who may respond best to treatment.

Developing rehabilitation treatments benefits from phased translational research as much or more than drug development, because the resources required for large-scale efficacy testing may be even greater in rehabilitation than in most pharmaceutical studies. Many rehabilitation treatments themselves are administered through lengthy interaction with a skilled rehabilitation therapist, with considerable labor cost. Accordingly, premature efficacy testing of rehabilitation treatments that have not been adequately evaluated, and may not be optimally designed, is wasteful. On the other hand, treatments that are systematically explored from a scientific perspective may not make their way to more definitive clinical testing. It is challenging but essential to coordinate the work of scientists across several levels of the translational continuum, so as to ensure that treatments move through this translational pipeline without getting stalled, while also ensuring that treatments are critically reframed and optimized at each developmental step.

TREATMENT THEORY AND ENABLEMENT THEORY

Treatment theory specifies the mechanism by which a proposed treatment changes its immediate treatment target—referred to here as the treatment object.^{5,14,15} Repetitive muscle contraction against a load (the treatment) leads to increased muscle strength (the treatment object) through a series of physiologic and biochemical events (the treatment mechanism). Treatment theory, however, is silent about the impact of this change in muscle strength on more distal aspects of function such as ambulation (here, distal refers to theoretical distance in the ICF framework, rather than anatomic/physical distance in its usual usage). Indeed, the impact of strengthening leg muscles on ambulation depends on a wide range of other coexisting abilities and impairments (eg, balance, visual perception), as well as environmental factors (eg, ground surface, lighting), and even social factors (safe home environment)—on which strengthening exercises have no impact.

Enablement theory, on the other hand, addresses the causal interrelationships among variables at different levels in the ICF.¹⁴ For example, enablement theory seeks to understand and specify the various capacities required for effective ambulation (eg, leg strength, balance, proprioception, visual perception), their relative importance to the activity of walking, and how they interact to determine walking ability, as illustrated in figure 2.

That is, enablement theory is primarily concerned with the causal arrows linking different entities in the ICF framework and, at a quantitative level, what the weightings of those causal links may be. Enablement theory, however, is silent about what mechanisms change leg strength as a variable contributing to walking; it only hypothesizes how distal activity variables will change if strength changes. In fact, enablement theory is completely unconcerned with which specific factors or treatments lead to the increased muscle strength, because its predictions depend only on the magnitude of change and the status of other relevant variables. Thus, treatment theory addresses mechanisms of an intervention—the traditional phase 0–III of pharmaceutical research—telling us how to change specific body structure and function variables, but does not address the functional impact of a change in an impairment. Enablement theory addresses the functional impact induced by various potential changes in impairments, but provides no tools for intervention.

Developing effective rehabilitation treatments that improve function requires an understanding of both treatment and enablement theory and their interface. If an ineffective treatment is applied, which fails to induce change in the treatment object, no improvement in more distal functional variables can be anticipated. On the other hand, a treatment may be mechanistically effective, and may improve the treatment object, but still fail to enhance more distal aspects of function without addressing other constraints on function, specified by enablement theory.

The distinction between treatment and enablement theories is relevant to translational rehabilitation research in 2 important ways. First, different phases of the treatment developmental trajectory rely more heavily on one or the other theoretical framework. Early phases of treatment development rely more heavily on treatment theory, because the goal is to develop a treatment that is maximally powerful in altering its desired object. For this purpose, outcome measures at the level of the treatment object are the most appropriate measures of efficacy. For example, in attempts to optimize muscle-strengthening exercises for the lower extremities, one is better off assessing efficacy with measures of strength than with measures of ambulation, because some patients may fail to improve in ambulation not because of failure of the strengthening exercises, but because untreated impairments in balance and proprioception continue to limit walking. A search for more effective strengthening exercises as a way of obtaining a treatment effect on walking would thus be misguided.

Second, because enablement theory is unconcerned with how a change in a variable occurred (but only that it did occur), there may not be a need for the functional impact of each new treatment that is focused on diminishing a particular impairment to be studied independently. That is, studies of the relationship between leg strength and walking ability (and balance and proprioception, etc) are relevant to all treatments that seek to enhance leg strength in the service of walking. Thus, we suggest that different groups of researchers, with greater expertise in either treatment theory or enablement theory, are likely to pursue different phases in this developmental trajectory. Successful translational rehabilitation research will require that these largely independent groups coordinate closely, or have effective methods to hand-off the research for further development. Specific interdisciplinary, organizational, and other administrative initiatives would be expected to enhance this collaboration.^{3,16}

PHASES OF TRANSLATIONAL REHABILITATION RESEARCH DEFINED

Despite the similar rationale for phased treatment research in pharmaceutical and rehabilitation research, the design challenges faced by rehabilitation research differ from pharmaceutical research in certain key ways. Phases of rehabilitation research are more

complex than pharmaceutical phase I, II, and III research, and the progress from idea inception to phase III and beyond is less linearly arranged. For this reason, in the next section we will focus primarily on the distinct goals and purposes of different research phases, rather than on their temporal sequence. Many rehabilitation treatments require iterative development. For example, one may develop a treatment algorithm or manual and measures of treatment adherence to capture the treatment's hypothesized active ingredients and verify their delivery in later phases of clinical research. But if early efficacy studies are disappointing, a researcher may be forced to return to the treatment development stage to try to intensify or modify the treatment ingredients, or improve the accuracy or specificity of the treatment object being measured in further testing. The following phases of rehabilitation research provide a smooth continuum of knowledge in a domain (see fig 1).

NATURAL HISTORY AND MEASUREMENT

Rigorous treatment research benefits from a thorough understanding of the problem to be treated, and the availability of good tools for measuring its manifestations and identifying variables associated with differences in outcome (ie, case-mix factors). There is definitely a need for more studies in rehabilitation that shed light on prevalence and natural recovery of a problem, to inform strategies for recruitment of research participants and the optimal timing for intervention studies. Identification of predictors of recovery, independent of the study treatment, may increase statistical power in later research phases, and even generate new treatment targets with the potential to influence the recovery trajectory. Identifying confounding factors and current treatments being delivered for the condition may constrain later study designs. Measuring the course of the problem with the same outcome measures intended for later studies also allows the estimation of measurement variance over time, and may highlight the need to develop new and more sensitive measurement tools before proceeding to treatment studies. For example, in laying the foundation for a large, multicenter randomized controlled trial (RCT) of amantadine for enhancing recovery from disorders of consciousness, Whyte et al¹⁷ conducted a natural history study of recovery in this population, using the primary outcome measure intended for that trial. Measurement of the variability in recovery trajectory allowed computation of the necessary sample size, and incorporation into the research design of prog-nostically predictive variables increased statistical power. Assessing enrollment of eligible patients by each center allowed assessment of study feasibility and selection of productive clinical sites. And assessment of the psychoactive treatments in clinical use in these centers defined the medical confounds that would need to be avoided or accommodated. Similarly, research on treating traumatic brain injury (TBI)-related attention deficits had to first focus on developing new measures of the clinical deficit, because few traditional attention measures reliably captured the attentional differences between patients and controls.¹⁸⁻²³

The issues of the clinical syndrome and its natural history, and measurement of the problem are, themselves, interrelated. That is, if one focuses on patients with attention deficits related to very severe TBI, one will be forced to choose observational measures of attention to capture change, because the patients cannot engage in standardized testing, whereas if one focuses on a patient group with milder deficits, various neuropsychologic tests and standardized work tasks may be more appropriate measurement tools.^{24,25} Moreover, if one intends to implement a treatment that spans this severity spectrum, either in its range of patients or in the range of their recovery trajectory, one must grapple with the measurement challenge that observational measures may be at ceiling for the milder group and neuropsychologic measures may be at floor in the more severe group. Thus, it is important that the outcome measure used is sensitive to the effects of treatment in the sample to be studied. Rehabilitation studies should also take into account confounds that can affect validity of outcome assessment tools. For example, unawareness of deficit can invalidate

self-report measures, and administering a horizontal Likert scale to stroke survivors with a rightward spatial orienting preference may bias reported values to the right of the scale, distorting measured values for quality of life or other patient-reported out-comes.²⁶

Such descriptive studies may also help clarify the link between impairments that are potential treatment targets and complex real-world function, in order to motivate a whole treatment direction. For example, authorities have questioned whether limb apraxia actually affects functional activities after stroke,²⁷ especially because stroke survivors and their families are rarely aware of the problem.²⁸ This might suggest that there is little point in developing treatments for this acquired disorder of skilled learned purposive movements. Foundas et al,²⁹ however, examined mealtime behaviors in stroke survivors with and without limb apraxia, as defined by semi-quantitative observational assessments. The study demonstrated that a group of left brain stroke survivors who made errors in the pathologic range on an impairment assessment sensitive to ideomotor limb apraxia (Florida Apraxia Battery³⁰) made more eating action errors during a meal, and completed fewer actions per unit time than those without the impairment.

IDEA GENERATION

As in pharmaceutical research, treatment development begins with the identification of a treatment of interest, and 1 or more potential treatment objects. Thus, idea generation operates at least covertly at the level of treatment theory (“This might be a useful treatment because ...”). Scientific interest in a potential rehabilitation treatment may emerge from animal models of recovery of function (eg, dextroamphetamine in animal motor recovery), tissue studies (eg, influences on axonal growth), studies of basic human cognition or motor physiology (eg, errorless learning in amnesia, properties of central pattern generators), or fortuitous observations of the effects of an existing treatment in another functional domain (eg, constraint-induced treatment transferred from paresis to aphasia). Regardless of the source of inspiration for this phase 0 research, sufficient interest in a treatment’s potential impact in humans will begin the translational process.

SAFETY AND DOSING

In pharmaceutical research, involving new drugs or use of older drugs in different patient groups, phase I is devoted to identifying the dose range that can be safely administered and the relevant side effects in the clinical population of interest, as well as pharmacokinetic factors that affect dosing. This allows early efficacy studies to be conducted with a minimum of risk but with optimal drug levels. Although safety assessment is formalized when pharmacologic or device treatments are investigated for their rehabilitation impact, it is often ignored when the treatments studied are behavioral or experiential in nature. However, it has been argued that any treatment that is hypothesized to be sufficiently potent to affect human function is also, in principle, capable of doing harm and, thus, should be assessed more rigorously for its safety.³¹ Determining where to look for safety risks in rehabilitation, however, may require creativity, because many of the adverse events of drugs emerge from their physiologic effects. In contrast, because the mechanisms of rehabilitation treatments are more varied, and some of their effects are functional in nature, adverse events could be as remote as psychological reactions to improvements or lack of improvements.⁵

Establishing an optimal dose of a rehabilitation intervention early in the research trajectory may be difficult. In many cases, adverse effects will not be the limiting factors to the maximal dose. Rather, issues such as cost, feasibility, and patient compliance are relevant. However, rehabilitation researchers must be wary of allowing economic aspects of current clinical practice, or existing expert clinical opinion, to overrule scientific prescription of study treatment doses. For example, constraint-induced movement therapy (CIMT) for

hemiparesis substituted 6 hours of treatment per day for 10 days over a 2-week period, for the traditional schedule of 1 to 2 hours of treatment, 3 times weekly, over several weeks to months, with improved effects.³² In the absence of efficacy data, it is also difficult to determine how costly or effortful is too costly or effortful. Thus, titration of treatment dosing may require iterative cycles of efficacy research and treatment revision.

PROOF OF CONCEPT

Research studies at this phase (analogous to phase II in pharmaceutical research) seek initial evidence that sheds light on the underlying mechanisms of a novel treatment phenomenon. In rehabilitation research, the phenomenon around which a report is structured is generally some form of improvement or recovery of function. Thus, the goal of such studies is to support the treatment theory by developing evidence that the experimental treatment modifies the treatment object as predicted, in a small number of relevant patients. The relevant treatment theory need not be biological. It can come from biomechanics, cognitive neuroscience, motor physiology, and many other scientific domains.

The important defining feature of a proof of concept study is that it does not attempt to produce data generalizable to an entire clinical group (ie, the treatment can work, rather than the treatment does work reliably). Because of this limitation, proof of concept case studies are properly followed with studies including a larger sample of subjects with a more representative range of symptoms or deficits. Indeed, a major purpose of proof of concept studies is to motivate the conduct of larger scale efficacy studies by demonstrating a treatment effect. Because minimizing type II error at this phase is more critical than minimizing type I error, matched control subjects receiving an intervention may not be required, and short-term biological responses to treatment may provide sufficient evidence to propel larger trials (eg, tumor regression after chemotherapy, prior to data on improved survival).

Proof of concept studies may be challenging in rehabilitation. Behavioral or experiential treatments must be well-defined in terms of their ingredients and administration before one can ask if they appear to have their desired effects. Thus, the laborious (and expensive) process of distilling these ingredients into some form of treatment manual, and developing measures to ensure the delivery of the ingredients, may need to occur before one can even obtain this preliminary assessment. In many rehabilitation studies, where treatments involve intensive behavioral experience to promote learning or plasticity, short-term assessment may provide little insight. Within-subject quasi-experimental designs with small samples (including well-controlled N-of-1 studies) may be used in some instances. Essentially, such experimental designs allow subjects to serve as their own controls, and may be useful when a treatment theory is extremely specific, requiring precisely characterized or uncommon deficits to be explored. These designs are also useful when subject heterogeneity limits the power of between-group comparisons, as when studying an impairment-directed treatment in stroke patients with aphasia who have a range of communication impairments (see Robey et al³³ for a review).

However, in the context of great behavioral variability and spontaneous recovery, even well designed N-of-1 studies may have difficulty quantifying the change induced by the treatment itself. Thus, the earliest proof of concept studies may require moderately large samples and a randomized controlled parallel group design to determine whether improvements seen are truly related to the treatment mechanism. Although randomized controlled parallel group designs are very useful, one must be cautious that designs that rely on a difference in means or medians between 2 treatment groups may not cope well with the heterogeneity among study patients. Hence, clinically relevant subgroup effects may mask

the activity of a treatment that is highly effective for some but ineffective for others, and lead such a treatment to fail at the proof of concept stage. For example, reversible recovery of consciousness in response to zolpidem administration unquestionably occurs after severe brain injury, but probably in less than 10% of unconscious patients.³⁴ Thus, a proof of concept parallel group RCT would likely conclude that this treatment is not worth pursuing. Because the clinician's interest is in individuals, and that of the literature is in groups, it is prudent to remember that comparing group means in any analysis obscures the largest effects that occur in some individuals. When a heterogeneous group of subjects is tested on multiple occasions over time, employing multilevel longitudinal modeling of random effects, otherwise disregarded in statistical analysis, can increase the ability to validly assess treatment-induced changes.³⁵

RCTs at the proof of concept stage are also often misinterpreted as phase III efficacy studies. As a result, grant or manuscript reviewers may judge these studies by harsher criteria than are appropriate for research at their developmental stage, expecting to see evidence of functional impact, quality of life change, or cost-effectiveness. Alternately, readers of published RCTs exploring proof of concept may not recognize that the study does not demonstrate efficacy or, certainly, effectiveness, and may prematurely begin the process to move treatment ideas into practice.

Proof of concept studies rely heavily on treatment theory and valid deductive reasoning. The research must thus demonstrate a detailed understanding of the object of the treatment (and how to measure its change with appropriate sensitivity), the clinical context (eg, likely patterns of change in the treatment object in the absence of the treatment), and the underlying mechanism of the treatment being applied (eg, what mental and physical capacities are required for a treatment response).¹⁵ Because treatments in these studies have had some initial feasibility assessment, but have not yet been evaluated for their ecological impact, treatment outcome is appropriately measured with respect to the treatment object alone. Ideally, a treatment would not move along the translational continuum until it's considered to be mechanistically optimized with respect to its effect on the treatment object. Then, researchers can investigate its more distal functional implications. For example, one may study the effects of methylphenidate (the treatment) on speed of information processing (the object) after TBI, with the aim of enhancing academic performance (an enablement outcome). Early research on this translational continuum will be concerned with whether methylphenidate does, indeed, speed information processing, and what dose maximizes this benefit without excessive side effects. Later trials, using this dose, can assess the generality of this effect across patient subgroups, the impact of speeding information processing on improving academic performance, and other outcomes.

EFFICACY STUDIES: CONTROLLING FOR CONFOUNDING INFLUENCES

As described under "Treatment Theory and Enablement Theory," a treatment theory specifies the treatment mechanism within a specific biological, psychological, chemical, kinesio-logical, or other framework. In other words, it specifies what chain of events, or interaction of physiological, cognitive, or other systems results in observable change in a treatment object. The treatment object may not have any impact on activities or participation, and may not even have any known relevance on body function (eg, grams of Huntingtin protein per 100 cells). However, efficacy studies are designed to evaluate the impact of the treatment on this treatment object, in a large and representative sample. To do so one must control for unrelated factors that may also alter the treatment object such as: spontaneous recovery (eg, after TBI, spinal cord injury, or stroke); practice effects or acquisition of testing set; and placebo effects—improvement or change in the treatment

object that results from awareness by the subjects that they are being treated or studied, rather than from the manipulation itself.

A major pitfall when carrying out efficacy studies is assuming that 1 set of solutions (using a placebo, randomly assigning subjects to a treatment or control condition) broadly addresses all internal and external validity problems to which the research might be vulnerable. An example of this kind of one size fits all solution is the use of placebos to control for confounding effects influencing improvement. Controlling for the influence of confounding factors using a placebo might seem to be a reasonable goal in drug studies, because, by definition, an inert pill posing as a drug may represent a fully plausible treatment while still being completely devoid of the treatment's active ingredients. In many areas of rehabilitation, however, a valid placebo treatment may not be available. If treatment includes voluntary activities such as exercises or repetitive practice, if the precise ingredients required for therapeutic impact of a complex therapy are not fully determined early in research (such that one can't be sure which of the ingredients are active and need to be controlled for), or if treatment involves devices whose physical properties are obvious to the participant, fully plausible and completely inactive may be mutually exclusive. In these studies, the experimenter may need to modify the study question so as to evaluate whether the treatment is preferable to natural evolution of the condition (eg, comparison with no treatment or a waitlist), a standard of care (which may be different in different institutions), other doses of the same treatment, or treatments that are hypothesized to work by different mechanisms. Even where placebos are feasible, numerous studies have demonstrated that the experimenter's knowledge or expectancy can induce a mock treatment effect,³⁶ requiring experimenter blinding even when participants receive a placebo.

Early efficacy studies typically select subjects, to the extent possible, who are most likely to respond to the treatment in a measurable way. That is, to the extent that the hypothesized treatment mechanism suggests who might be responders and nonresponders, inclusion and exclusion criteria for clinical trials are generally designed to maximize response rate. For example, in a recent study of a treatment to enhance recovery from the vegetative state, we limited enrollment to 16 weeks postinjury on the assumption that those subjects who had been unconscious for longer intervals might have brain damage so severe that few treatments would benefit them.³⁷ Similarly, many studies of CIMT for upper extremity paresis have required some finger extension as a condition of participation, and don't enroll subjects with more severe motor impairment.³⁸ Because they are performed in a group selected to represent potential responders, efficacy studies may overestimate the impact of treatment on a broader population. Additional efficacy or effectiveness studies need to study response among patient subgroups that are hypothesized to respond differently to the treatment. A well-developed treatment theory may help define these subgroups. For example, if a compensatory cognitive strategy requires explicit memory for its application, this may suggest that amnesic patients are unlikely to benefit. However, subgroups also emerge from appropriate data analysis of samples in later efficacy and effectiveness studies.

Because definitive efficacy research typically requires relatively large samples, studies at this stage are often multicenter in nature. Consequently, they must deal with the possibility that clinical characteristics of patients and/or the quality of standard care vary systematically by site. For this reason, it is important that treatment assignment be stratified by site and often that data analysis be conducted in a nested fashion in which treatment site, as well as the treatment effect, are explicitly modeled.³⁹

EFFECTIVENESS RESEARCH—GENERALIZABILITY OF TREATMENT EFFECTS

In biomedical research, a distinction is often made between efficacy and effectiveness, with effectiveness more closely reflecting the impact of a treatment in routine clinical practice. Note that the goals of efficacy and effectiveness studies also differ. Efficacy research is typically still designed with the treatment mechanism in mind—that is, with selecting participants and outcome measures to provide maximal evidence on the success of the hypothesized treatment mechanism. In contrast, effectiveness studies are more focused on practical and policy questions. Given that this treatment has been shown to have a particular mechanism of action, in whom does that mechanism of action provide substantial practical benefits at a minimal risk and cost? As noted, typical efficacy trials involve carefully screened patients who are judged motivated and compliant, and clinical investigators who are well-trained and closely supervised. However, once a treatment is approved, it is likely to be administered to patients who are much more heterogeneous, and by clinicians of varied levels of experience. Thus, a valid concern of effectiveness research is that the degree of benefit seen in efficacy studies may not be replicated in community application. This same concern exists in rehabilitation studies. Will the typical patients with hemiparesis seen in multispecialty hospital settings be as motivated to participate in CIMT as those who signed up for the experimental study? In rehabilitation, the gap between efficacy and effectiveness studies may be subdivided along the lines of previously discussed treatment versus enablement theories.

A more limited form of rehabilitation effectiveness assessment retains the focus on the treatment object, and is very similar to effectiveness research in the medical arena. That is, when the treatment is delivered to more varied patients by more varied professionals, does it still have a potent effect on the treatment object? One might ask whether improvements of motor control of the hand are achieved in CIMT, when applied to similar patients in community hospitals or to patients with more severe motor impairment. If the answer is no, then one must search for the factors that account for the reduced impact. One might, for example, need to refine the treatment manual to be more supportive of clinicians with less sophisticated backgrounds, or to develop quick screening tests that can help identify the patients most likely to benefit from the treatment. Or one may discover feasibility constraints that require different solutions, as when a clinic-based treatment encounters patient transportation problems that interfere with compliance, leading either to transportation assistance or telemedicine solutions. One may even, at this late stage, entertain some revision of the active ingredients of the treatment, as when, for example, it turns out that clinicians who present the treatment with specific goal-setting maneuvers achieve better treatment responses, and structured goal setting is incorporated into the treatment. Ultimately, the job of this form of effectiveness research is to document good effectiveness, to find ways to enhance the treatment capacities of average clinicians, or to find ways to enhance the treatment response (defined by the treatment object) among poor responders.

EFFECTIVENESS RESEARCH—IMPACT ON ACTIVITY AND PARTICIPATION OUTCOMES

The concept of effectiveness in rehabilitation typically implies a change in meaningful function—that is, improved ability to engage in important life activities and societal participation. Consequently, evidence for effectiveness in this sense must go beyond a demonstration of the widespread ability to influence a specific impairment. Addressing this research goal requires attention to corresponding enablement theories. When clinicians

select a treatment to enhance a patient's activity or participation goals, a key challenge is selecting patients where this impact can be achieved. As noted, this problem cannot be solved by strengthening the treatment. Rather, it must be solved by studying the impact of combinations of treatments in patients who have multiple obstacles in the way of their activity goals, and/or limiting treatment recipients to the set of patients whose activity or participation goals are predominantly limited by the treated impairment.

For example, consider attention process training (APT), one of the most frequently used behavioral methods for remediating attention deficits related to TBI.⁴⁰ One might seek to enhance attentional function as a means of helping patients with TBI return to work in a productive manner. Even if APT provides reliable improvement in specific attentional processes, its impact on employment may be quite variable because some patients with poor attention also have inappropriate social behaviors and some also have severe motor dysfunction. One may need to define a subgroup of TBI patients with good social and motor function, but impaired attention, who are predicted to improve their vocational prospects from APT alone. Alternatively, one might develop an integrated treatment milieu which offers APT, movement rehabilitation, and training in social behavior skills, and assess the impact of this aggregate treatment on vocational outcome. The ultimate goal of this form of effectiveness research is to define the group of patients who will experience meaningful functional benefit from this treatment alone, or to define a group of patients who will benefit from a combined therapy program that includes this treatment. Research that stops short of this step leaves individual clinicians to do the complex enablement predictions without benefit of evidence.

As noted previously, however, evidence pertaining to this question can accrue from separately conducted research, which doesn't focus specifically on APT or any particular treatment for attention deficits. One can study, more generally, the relative contributions of attention skills, social behavior, and motor function to job success. If one finds that attention abilities appear predictive of work performance primarily in patients with good social and motor function, this will provide guidance to treating clinicians about how to apply APT and a range of other treatments that may have similar effects on attention.

We have argued that one may find a treatment to be efficacious and even effective in modifying its treatment object (ie, evidence has supported the treatment theory) without yet understanding the role of the treated impairment in larger activity and participation goals (the enablement theory). One may also examine the interrelationships among impairments in determining a functional capacity (ie, evidence supports an enablement theory) without yet having an effective treatment for a key impairment (ie, treatment theory not yet supported). This has implications for how we conceptualize the trajectory of translation. Note that in figure 1, we have placed the step of effectiveness translation from impairment to activity/participation (obstacle #3) after the step of clinical adoption of the treatment (obstacle #2). In such a model, clinicians may begin to adopt treatments that are effective in ameliorating impairments but have little practical value on activity/participation. Thus, one might argue that this effectiveness research step should precede clinical adoption. However, this would require that each treatment be studied separately with respect to its place in an enablement model, to ensure its impact on functional outcomes. Because, as we have argued, these 2 forms of research may take place relatively independently by different research groups and achieve efficiencies in doing so, it may be more realistic for clinicians to be armed with a growing set of treatment tools effective in ameliorating impairments. Their choice of which of those tools to employ with which patients to achieve real-world benefits will continue to evolve over time in response to developments in enablement research. Responsible clinicians will also make controlled observations of effects of impairment treatments on

activity and participation on an individual basis, in order to judge the success of the treatment algorithm for a particular patient under their care.

Consider research on the treatment of spasticity as a case in point. Over the last several decades, a number of oral and injectable medications, surgical interventions, and physical modalities have been shown to reduce spasticity as their direct treatment object. Research has also evolved showing that the relevance of spasticity as an obstacle to function varies greatly from patient to patient, depending on the degree of preserved strength, coordination, and range of motion. Thus, clinicians continue to evolve in their selection of patients for application of these treatments, and in their selection of other treatments to apply concurrently (eg, strengthening, serial casting), but this evolving enablement modeling alters the role of all of these treatments in achieving clinical effectiveness.

Effectiveness research has received increased attention in recent years, based on the claim that traditional RCTs have not provided optimal guidance to practicing clinicians and policy makers. Tunis et al¹⁶ have suggested the practical clinical trial in which realistically heterogeneous patients from community sites are randomized to available treatments which practicing clinicians must choose between. Such trials typically collect more streamlined data on participants and outcomes to constrain costs while obtaining an answer of clinical relevance. This kind of trial may be useful for exploring the first sense of effectiveness as described under “Effectiveness Research— Generalizability of Treatment Effects” (ie, widespread efficacy in addressing the treatment object), but may be more difficult to apply in service of enablement modeling. For the latter purpose, it is helpful to have a priori predictions of what impairments are relevant to the model, and to measure them as well as the treated impairment. Moreover, as noted before, informative enablement models can be developed outside the context of studies of 1 or 2 specific treatments.

HEALTH SERVICES RESEARCH

The most mature phase of rehabilitation research assesses patterns of practice and real-world outcomes in the community. Even effectiveness research, as discussed above, is conducted on patients whose treatment is selected by the study design. But do practicing clinicians choose this treatment for appropriate patients? Are there regional variations in the availability or quality of the treatment? What does it cost when routinely delivered?

Health services research usually relies on routine clinical and administrative data sources. Unfortunately, in the case of behavioral or experiential rehabilitation treatments, there is to date no operationally defined taxonomy by which such treatments are coded. Indeed, most rehabilitation treatments are coded, for administrative purposes, either in terms of the discipline of the clinician providing them (speech therapy), or the intended outcome of the treatment (eg, memory remediation), along with the time devoted, rather than by any system that corresponds to their active ingredients. Even the more nuanced coding systems used in recent practice-based evidence studies,⁴¹ bear an unknown correspondence to the active ingredients of treatment. Thus, at present, health services research in rehabilitation must confine itself to gross measures of service type (eg, hospital vs nursing home-based rehabilitation), amount (eg, hours of physical therapy), or goal-related label (eg, gait training). As more theory-based and manualized treatments are developed, however, it will become increasingly possible to evaluate the effects of rehabilitation treatments delivered in routine clinical care.

CONCLUSIONS

In this article, we have addressed the importance of a phased developmental approach to translational treatment research in general. We have also described its critical importance in

rehabilitation treatment research—particularly research on treatments intended to ameliorate specific impairments—and some key differences from the research phases designed for pharmaceutical research development. This phased development framework is intended to increase the likelihood of obtaining internally and externally valid evidence of treatment effects, thus directing investment of resources toward the most promising treatments. By combining the insights of treatment and enablement theories, we can ensure that the potency of treatments is optimized during development, and that the patients who will experience the greatest impact on activities and participation are identified. Additionally, this approach ensures that enough is known about the treatments to be tested that even negative study results will help to refine and advance the research program, thus building rehabilitation science in addition to adding empirical evidence of treatment efficacy.

Acknowledgments

Supported in part by the National Institutes of Health (grant nos. HD050836 and HD062647) and National Institute on Disability and Rehabilitation Research (cooperative agreement no. H133A080053).

List of Abbreviations

APT	attention process training
CIMT	constraint-induced movement therapy
ICF	<i>International Classification of Functioning, Disability and Health</i>
RCT	randomized controlled trial
TBI	traumatic brain injury

References

1. Contopoulos-Ioannidis DG, Ntzani EE, Ioannidis JP. Translation of highly promising basic science research into clinical applications. *Am J Med.* 2003; 114:474–484.
2. Woolf SH. The meaning of translational research and why it matters. *JAMA.* 2008; 299:211–213. [PubMed: 18182604]
3. Sung NS, Crowley WF Jr, Genel M. Central challenges facing the national clinical research enterprise. *JAMA.* 2003; 289:1278–1287. [PubMed: 12633190]
4. Whyte J, Gordon W, Rothi LJ. A phased developmental approach to neurorehabilitation research: the science of knowledge building. *Arch Phys Med Rehab.* 2009; 90(11 Suppl):S3–S10.
5. Whyte J. Directions in brain injury research: from concept to clinical implementation. *Neuropsychol Rehabil.* 2009; 19:807–823. [PubMed: 19548176]
6. Barrett AM, Levy CE, Rothi LJ. Post-stroke and brain injury rehabilitation: treatment strategies. *Am J Phys Med Rehab.* 2007; 86:694–695.
7. Rothi LJ. Cognitive rehabilitation: the role of theoretical rationales and respect for the maturational process needed for our evidence. *J Head Trauma Rehab.* 2006; 21:194–197.
8. Stucki G. International Classification of Functioning, Disability, and Health (ICF): a promising framework and classification for rehabilitation medicine. *Am J Phys Med Rehab.* 2005; 84:733–740.
9. Kummar S, Rubinstein L, Kinders R, et al. Phase 0 clinical trials: conceptions and misconceptions. *Cancer J.* 2008; 14:133–137. [PubMed: 18536551]
10. Rothi LJ, Barrett AM. The changing view of neurorehabilitation: a new era of optimism. *J Int Neuropsych Soc.* 2006; 12:812–815.
11. Guyatt G, Sackett D, Taylor DW, Chong J, Roberts R, Pugsley S. Determining optimal therapy—randomized trials in individual patients. *New Engl J Med.* 1986; 314:889–892. [PubMed: 2936958]

12. Edlund, W.; Gronseth, G.; So, Y.; Franklin, G. Clinical practice guideline process manual. St. Paul: American Academy of Neurology; 2004.
13. French J, Gronseth G. Lost in a jungle of evidence: we need a compass. *Neurology*. 2008; 71:1634–1638. [PubMed: 19001254]
14. Whyte J. A grand unified theory of rehabilitation (we wish!). The 57th John Stanley Coulter Memorial Lecture. *Arch Phys Med Rehab*. 2008; 89:203–209.
15. Whyte J. Using treatment theory to refine the designs of brain injury rehabilitation treatment effectiveness studies. *J Head Trauma Rehab*. 2006; 21:99–106.
16. Tunis SR, Stryer D, Clancy CM. Practical clinical trials: increasing the value of clinical research for decision making in clinical and health policy. *JAMA*. 2003; 290:1624–1632. [PubMed: 14506122]
17. Whyte J, Katz D, Long D, et al. Predictors of outcome and effect of psychoactive medications in prolonged posttraumatic disorders of consciousness: A multicenter study. *Arch Phys Med Rehab*. 2005; 86:453–462.
18. Whyte J, Fleming M, Polansky M, Cavallucci C. The effects of visual distraction following traumatic brain injury. *J Int Neuropsych Soc*. 1998; 4:126–136.
19. Whyte J, Fleming M, Polansky M, Cavallucci C, Coslett HB. Phasic arousal in response to auditory warnings after traumatic brain injury. *Neuropsychologia*. 1997; 35:313–324. [PubMed: 9051679]
20. Whyte J, Grieb-Neff P, Gantz C, Polansky M. Measuring sustained attention after traumatic brain injury: differences in key findings from the sustained attention to response task (SART). *Neuropsychologia*. 2006; 44:2007–2014. [PubMed: 16682059]
21. Whyte J, Polansky M, Cavallucci C, Fleming M, Lhulier J, Coslett H. Inattentive behavior after traumatic brain injury. *J Int Neuropsych Soc*. 1996; 2:274–281.
22. Whyte J, Polansky M, Fleming M, Coslett H, Cavallucci C. Sustained arousal and attention after traumatic brain injury. *Neuropsychologia*. 1995; 33:797–813. [PubMed: 7477808]
23. Whyte J, Schuster K, Polansky M, Adams J, Coslett HB. Frequency and duration of inattentive behavior after traumatic brain injury: effects of distraction, task and practice. *J Int Neuropsych Soc*. 2000; 6:1–11.
24. Whyte J, Hart T, Bode R, Malec JF. The Moss Attention Rating Scale for traumatic brain injury: initial psychometric assessment. *Arch Phys Med Rehab*. 2003; 84:268–276.
25. Whyte J, Hart T, Ellis CA, Chervoneva I. The Moss Attention Rating Scale for traumatic brain injury: further explorations of reliability and sensitivity to change. *Arch Phys Med Rehab*. 2008; 89:966–973.
26. Barrett AM. Rose-colored answers: neuropsychological deficits and patient-reported outcomes after stroke. *Behav Neurol*. 2010; 221:17–23. [PubMed: 20543455]
27. Poeck, K. Clues to the nature of disruptions to limb praxis. In: Roy, EA., editor. *Neuropsychological studies of apraxia and related disorders*. New York: North-Holland; 1985. p. 99-107.
28. Barrett, AM.; Foundas, AL. Apraxia. In: Rizzo, M.; Eslinger, PJ., editors. *Principles and practice of behavioral neurology and neuropsychology*. Philadelphia: Saunders/Churchill/Livingstone/Mosby; 2004. p. 409-422.
29. Foundas AL, Macauley BL, Raymer AM, Maher L, Heilman KM, Rothi LJ. Ecological implications of limb apraxia: evidence from mealtime behavior. *J Int Neuropsych Soc*. 1995; 1:62–66.
30. Rothi LJG, Raymer AM, Ochipa C, Maher LM, Greenwald ML, Heilman KM. *Florida Apraxia Battery*, experimental edition. 1992
31. Barrett AM, Schwartz RL, Crucian GP, Heilman KM. Adverse effect of dopamine agonist therapy in a patient with motor-intentional neglect. *Arch Phys Med Rehab*. 1999; 80:600–603.
32. Taub E, Miller NE, Novack T, et al. Technique to improve chronic motor deficit after stroke. *Arch Phys Med Rehab*. 1993; 74:347–354.
33. Robey RR, Schultz MC, Crawford AB, Sinner C. Single subject clinical outcome research: designs, data, effect sizes and analyses. *Aphasiology*. 1999; 13:445–473.

34. Whyte J, Myers R. Incidence of clinically significant responses to zolpidem among patients with disorders of consciousness: a preliminary placebo controlled trial. *Am J Phys Med Rehab.* 2009; 88:410–418.
35. Kwok OM, Underhill AT, Berry JW, Luo W, Elliott TR, Yoon M. Analyzing longitudinal data with multilevel models: an example with individuals living with lower extremity intra-articular fractures. *Rehabil Psychol.* 2008; 53:370–386. [PubMed: 19649151]
36. Rosenthal, R. Interpersonal expectations: effects of the experimenter's hypothesis. In: Rosenthal, R.; Rosnow, RL., editors. *Artifacts in behavioral research.* New York: Oxford Univ Pr.; 2009. p. 138-210.
37. Whyte J. Treatments to enhance recovery from the vegetative and minimally conscious states: ethical issues surrounding efficacy studies. *Am J Phys Med Rehab.* 2007; 86:86–92.
38. Wolf SL, Winstein CJ, Miller JP, et al. Effect of constraint-induced movement therapy on upper extremity function 3 to 9 months after stroke: the EXCITE randomized clinical trial. *JAMA.* 2006; 296:2095–2104. [PubMed: 17077374]
39. Raudenbush, SW.; Bryk, AS. 2nd ed.. Thousand Oaks: Sage Publications; 2002. *Hierarchical linear models: applications and data analysis methods.*
40. Sohlberg MM, McLaughlin KA, Pavese A, Heidrich A, Posner M. Evaluation of attention process training and brain injury education in persons with acquired brain injury. *J Clin Exp Neuropsych.* 2000; 22:656–676.
41. Horn SD, Gassaway J. Practice based evidence: incorporating clinical heterogeneity and patient-reported outcomes for comparative effectiveness research. *Med Care.* 2010; 48(6 Suppl):S17–S22. [PubMed: 20421825]

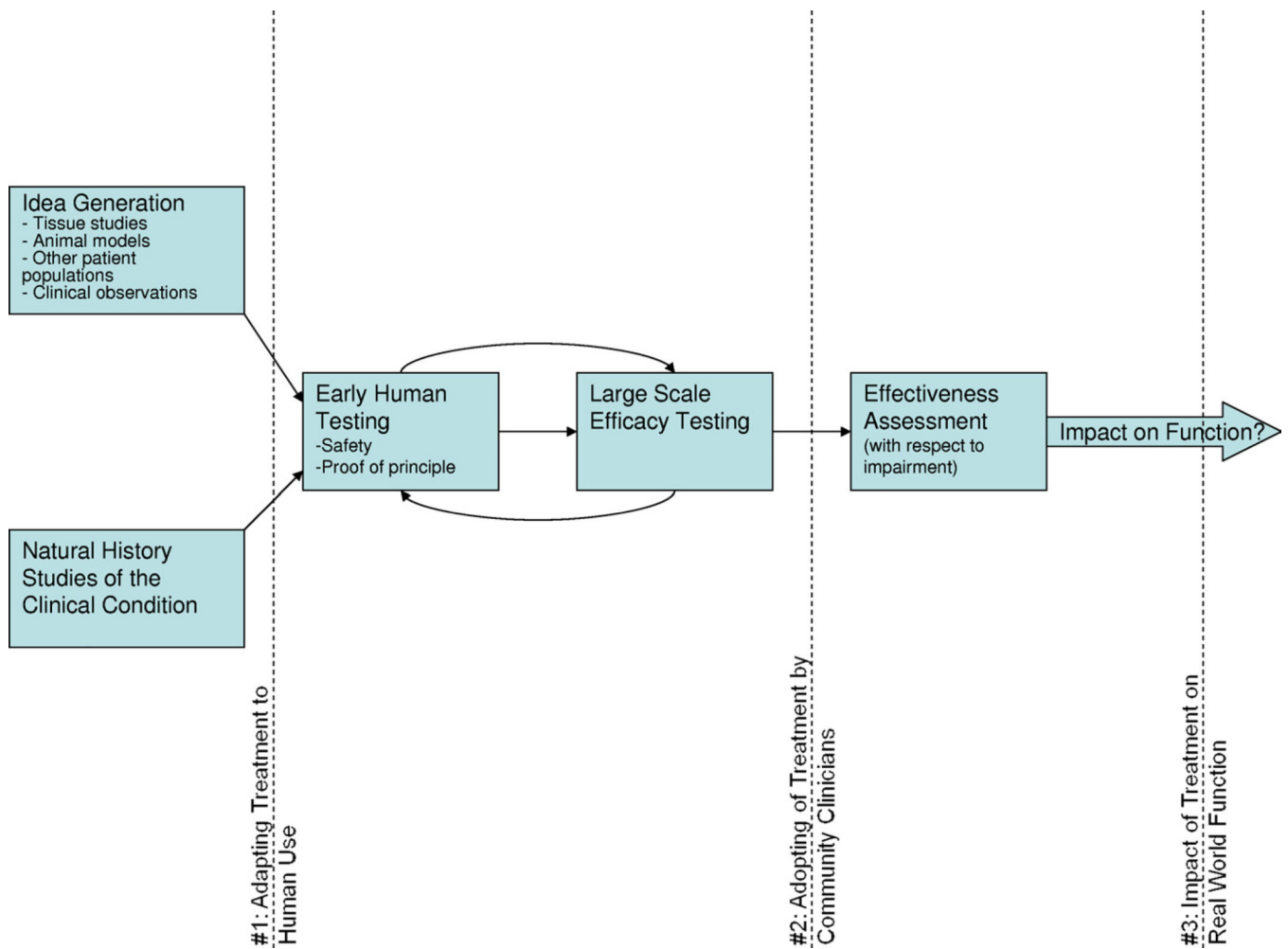


Fig 1. Shows the general trajectory of translational rehabilitation research. Early human testing benefits from a thorough knowledge of the natural history of the clinical condition, coupled with a novel treatment idea, which may emerge from a variety of sources. The initial translational bottleneck (dotted line #1) involves the challenge of adapting the treatment concept to feasible human implementation. Human safety, feasibility, and proof of principle studies ultimately lead to large-scale efficacy testing, though multiple iterations may be required. Once efficacy is established, the next translational obstacle (dotted line #2) involves successful dissemination of the new treatment to practicing clinicians in the community. Substantial adoption paves the way for exploration of effectiveness in the hands of these typical clinicians and patients, still focused on impairment-level outcomes. Translational bottleneck #3 involves the question of real-world impact of the treatment on more macro outcomes, primarily dependent on enablement theory (see fig 2). Adapted from Sung et al.³

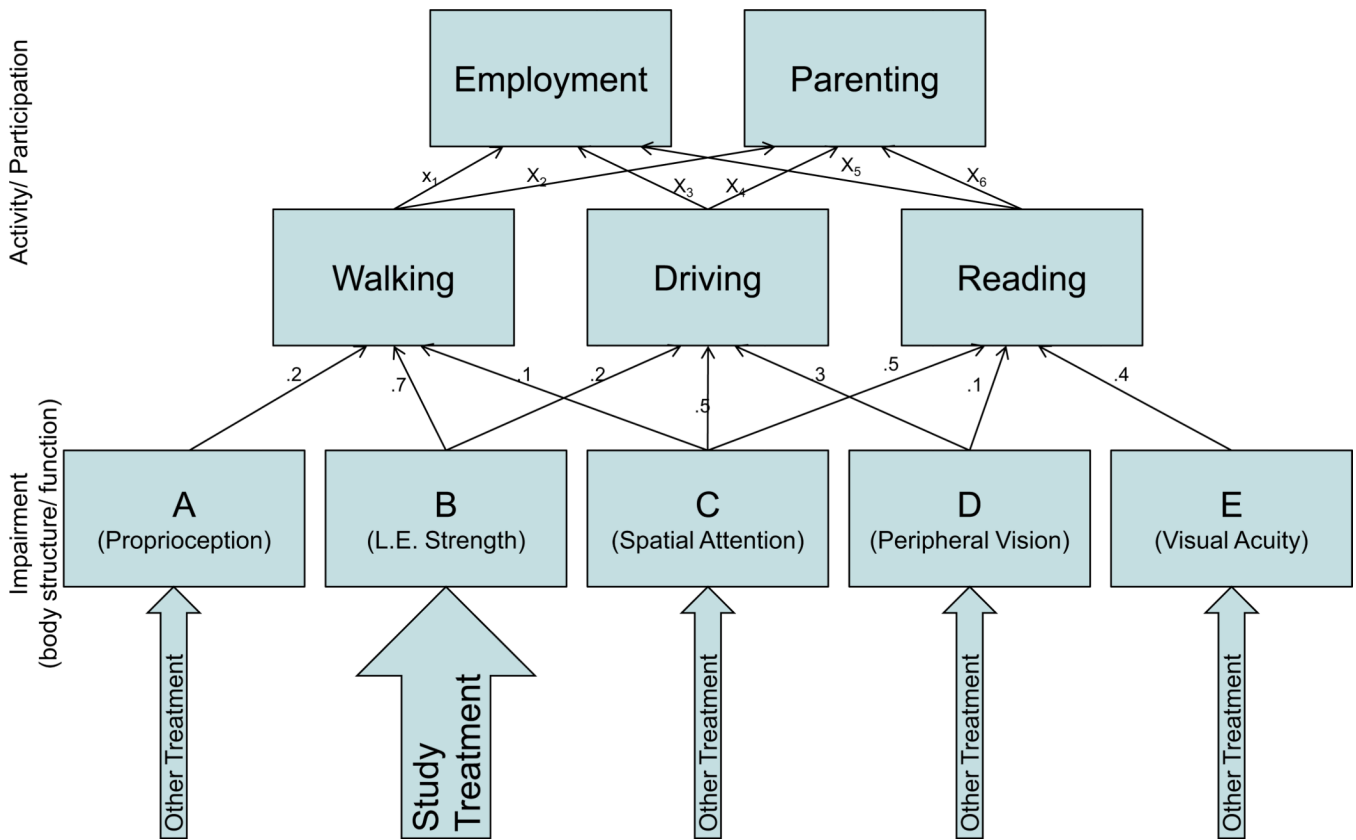


Fig 2. Shows how one of the authors (J.W.) would illustrate the hypothetical relationship among variables in the ICF framework, at impairment (body structure/function) and activity/participation levels. Note that several variables at a lower level contribute to performance at the next level (a highly oversimplified model in this example, in that each activity depends on only 3 body functions), with the numbers next to each arrow indicating hypothetical weights reflecting the importance of each variable in determining performance at the next level. Thus, for example, the activity of walking is strongly determined by performance capacity B (lower extremity strength), but less strongly by performance capacities A and C (proprioception and spatial attention, respectively). In contrast, the activity of driving is less strongly dependent on capacity B (lower extremity strength) but has a stronger relationship to capacities C (spatial attention) and D (peripheral vision). The enablement theory involves hypothesizing and subsequently determining the locations of the relevant causal arrows in such a schema and their relative weights. Note several implications of this model: (1) in general, improvement in capacity B (lower extremity strength) would be expected to have greater impact on walking than on driving; (2) isolated application of the study treatment, even if effective, would be predicted to lead to substantial improvements in walking for many patients, but in driving only for those patients with spatial attention and peripheral vision (ie, studies seeking to demonstrate impact of the study treatment alone on driving should select patients with normal spatial attention and peripheral vision); and (3) in a heterogeneous patient population, the study treatment may need to be combined with treatments for capacities C and D in order to see a widespread impact on driving.

Table 1

Phases of Translational Research

Stage of Rehabilitation Research Development	Correspondence to Pharmaceutical Phased Research Development	Translational Task and Obstacles
Idea inception	Phase 0 ⁹	Viewing unexpected observations/findings and negative results as opportunity to generate ideas rather than as trivia/research errors is critical to innovative idea generation.
Natural history and measurement	Contributes to phase I—the underlying recovery trajectory and object of treatment needs to be delineated in order to detect adverse effects and treatment effects	Relationship of treatment mechanism to object of treatment under study must be defined (treatment theory). Relationship of object of treatment and outcome measures must also be clear.
Proof of concept	Phase I, Phase II	Refinement of treatment theory may radically alter the idea under study. Paradigmatic rigidity (no useful stroke recovery can occur after 6mo) could block further development.
Evaluate efficacy	Phase III	Define experimental conditions so as to limit variability/ confounds while avoiding type II error and generating feasible protocols for the next study phase. New statistical analytic techniques (modeling) may be highly useful.
Evaluate effectiveness	Phase IV	Requires rigorous implementation of enablement theory, even when treatment theory has guided process up to this point.