

Computerized behavioral-data collection and analysis for improved clinical outcomes in rehabilitation

CHARLES T. MERBITZ

*Illinois Institute of Technology, Rehabilitation Institute of Chicago
and Northwestern University Medical School, Chicago, Illinois*

ROSEMARIE B. KING

Rehabilitation Institute of Chicago, Chicago, Illinois

LEORA R. CHERNEY

*Rehabilitation Institute of Chicago and Northwestern University Medical School
Chicago, Illinois*

HANSPETER MARQUI

Rehabilitation Institute of Chicago, Chicago, Illinois

JEFFREY C. GRIP

*Rehabilitation Institute of Chicago and Chicago School of Professional Psychology
Chicago, Illinois*

and

TRACY J. MARKOWITZ

Illinois Institute of Technology, Chicago, Illinois

Medical rehabilitation seeks change in patient behavior. Because of the variability in behavior across individuals and the uncertainty of clinical outcomes, an appropriate model for clinical intervention across the rehabilitation disciplines is single-case design. Therefore, the rehabilitation process can be facilitated by data collection and analysis tools similar to those used in the experimental analysis of behavior. We describe a system of computerized tools for the measurement and control of parts of this clinical process.

Increasing resources are devoted to the rehabilitation of persons with physical disabilities. Etiology, loci, and extent of impairment vary, and many persons have multiple concurrent medical problems. Severity ranges from mild to very severe.

Medical-rehabilitation professionals seek to improve the functioning and prevent further impairment of these persons. Since "improved functioning" means "improved behavior" (Kaplan, 1990), "learning" is a key construct in rehabilitation treatment and a large proportion of effort goes to changing behavior (Bleiberg & Merbitz, 1983). Payers (government, insurance companies, and individuals), professionals, families, and patients alike desire cost-effective treatments that induce maximum possible

appropriate sustained behavior change for each individual. However, since many persons in rehabilitation have damage to the brain and central nervous system, impairments in learning are not considered unusual. Similarly, patients exhibit the range of personality variables found in the general population, possibly exacerbated in many cases by other variables, such as medication side effects, brain damage, the circumstances of suddenly being treated as dependent, labile, and irresponsible, and so forth. Thus, a broad range of goals and interventions is required to meet the breadth of patient needs, and fitting interventions to patients is an ongoing challenge.

Areas of deficit for an individual are delineated by diagnostic assessment, and then interventions are prescribed. The large individual differences between patients cited above have impeded the use of traditional group-oriented statistical designs in the development and validation of effective interventions, and single-subject designs are often used to support claims of effectiveness of interventions.

Single-subject designs are also recommended as the model for daily management of behavior-change processes

Partial support for the development of these systems was provided by the National Institute for Disability and Rehabilitation Research through Grants G00830079 and H133B80007 and by the Rehabilitation Institute Foundation of the Rehabilitation Institute of Chicago. Correspondence should be addressed to C. T. Merbitz, Department of Psychology, Illinois Institute of Technology, IIT Center, Chicago, IL 60616.

in rehabilitation, since there are few powerful interventions with precisely known desirable outcomes. A medical model of such designs is the system of empirical medicine (Bernard, 1865/1957). The basic principle is that of experimental analysis (frequently measure the behavior and modify the treatment according to its success). In rehabilitation nursing and the allied health disciplines such as occupational therapy, physical therapy, psychology, and speech-language pathology, the interventions (from which the patients are to learn better performance) are themselves often composed of behaviors emitted by the clinician. Thus, the clinician ideally provides a behavioral environment within which the patient learns at an optimum rate; often multiple clinical goals are pursued.

From the point of view of the individual professional, effective deployment of this system requires measures that can distinguish between improved, unchanged, or deteriorated functioning for each targeted behavior. Also, as the treatments are to be finely adjusted, the measures must be amenable to frequent reapplication. Logically, the system rests on projections of the expected performance of each behavior if treatment is unchanged, so measures should also provide for predictions. These are stringent requirements. Deploying measurement and graphing for experimental analysis of five or six behaviors per patient per discipline would substantially increase the cost of services.

While such graphing is used in some situations, for ongoing management "clinical judgment" is more common, supplemented by periodic testing and informally developed measures. Also, rehabilitation organizations have developed some nominal and ordinal measures of general functioning for program evaluation; however, these are of little use in managing the treatment of individuals (Merbitz, Morris, & Grip, 1989).

PRECISION REHABILITATION

The Precision Rehabilitation System (PRS) is a micro-computer-based system for inexpensive, data-based, feedback-driven management of the rehabilitation process. Data in this system are recorded and reported as events in time (Johnston & Pennypacker, 1980), yielding ratio-quality measurement across a variety of behaviors. Typically, both patient and clinician behaviors are recorded so that the clinician can examine his/her implementation of the planned intervention and correct flaws or otherwise modify the approach when progress is not satisfactory. Thus, the system resembles a laboratory data collection and analysis system optimized for single-subject designs and the clinical setting.

Figure 1 shows the system design. The PRS runs currently on the Apple IIe, IIc, or IIgs under ProDOS. To provide good performance, the PRS was programmed in assembly language. Applications in PRS have three parts: a generic data-analysis module (DADS, or Data Analysis/Decision Support), a data-file management module (MOMS, or Master Organizer, Manager, and Saver), and an application module. The application module and MOMS are

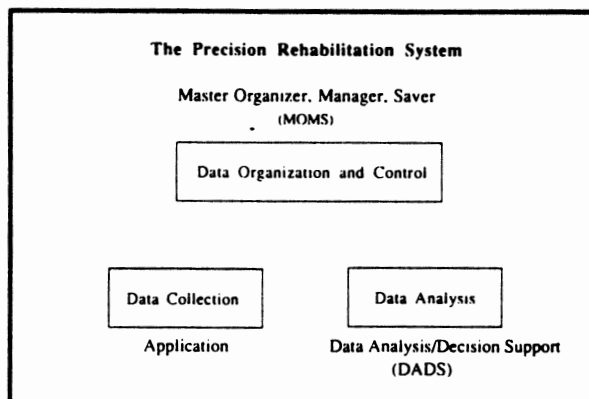


Figure 1. Overall system design.

customized for the specific data-collection and stimulus-display requirements of each clinical situation. The former collects data and provides stimuli during data collection; MOMS is used off line to set up for the session. Each clinical session is saved as a separate disk file; DADS is then invoked. Several clinical applications are discussed below, followed by brief descriptions of data-file structure and DADS.

Applications

Earlier versions of some system applications have been demonstrated elsewhere. We here describe two of the more mature applications in some detail and briefly mention two more.

The CAS (Communication Analysis System) provides ongoing measures of 21 client and clinician communicative behaviors during speech-language treatment (Merbitz, Cherney, & Marqui, 1991, 1992; Merbitz, Grip, et al., 1989). During the clinical session, the CAS runs in an Apple and the clinician presses keys to record the occurrence of critical behaviors. Figure 2 shows data from a person with aphasia performing an oral reading exercise. The cumulative record shows words correctly read. The record resets with each new sentence. The event records show emissions of two classes of error responses. Figure 3 shows the change in frequencies of correct and incorrect emissions over 4 months of rehabilitation treatment.

Another application, the TLC (Timer-Logger-Communicator) is a battery-powered CMOS single-board 32K microcomputer that provides remote 24-h mobile recording of behavior for persons in wheelchairs (Grip & Merbitz, 1986; Merbitz, King, Bleiberg, & Grip, 1985). In rehabilitation, a goal is to teach patients with spinal-cord injury to emit a pressure-relief behavior (PRB) once every 20 min when seated in the wheelchair. Such behavior is thought to prevent expensive and health-threatening pressure sores. Figure 4 shows frequency (count per minute) of PRBs over the course of hospitalization for one patient. Figure 5 is a plotter-drawn graphic designed for clinical

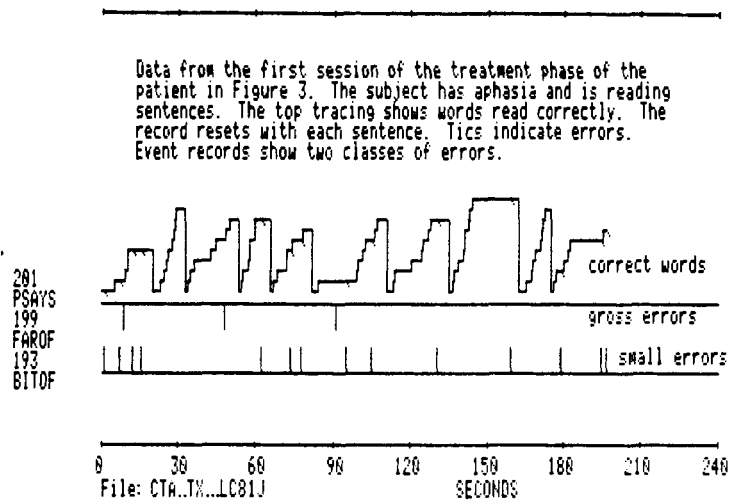


Figure 2. Sample cumulative and event records generated by the system. The cumulative record reset is controllable. Phase lines and comments can be placed anywhere on the screen. The file name is listed at the bottom of the figure, and the event numbers and labels are at left.

use that shows a single day of wheelchair use and the patient's PRBs. Patients and staff discuss the data each day to support appropriate performance.

Other applications that have been tested provide simplified data recording for clinical analysis of gait rehabilitation progress (Roth et al., 1990) and cognitive retraining games for training persons with brain injury.

Accuracy, Validity, and Reliability

Accuracy (Johnston & Pennypacker, 1980), validity, and reliability are addressed separately for each applica-

tion, but some commonalities apply across the system. Validity is generally not a major issue, since system applications are not designed to measure internal psychological constructs. Instead, they measure overt behaviors that are themselves important and whose change is desirable.

Reliability is usually not a major theoretical problem of the system as long as transduction of the behavior remains reliable, and physical reliability of the commercially purchased elements of the system has been excellent. However, the TLC and gait applications are implemented on a custom board, and physical reliability of these elements

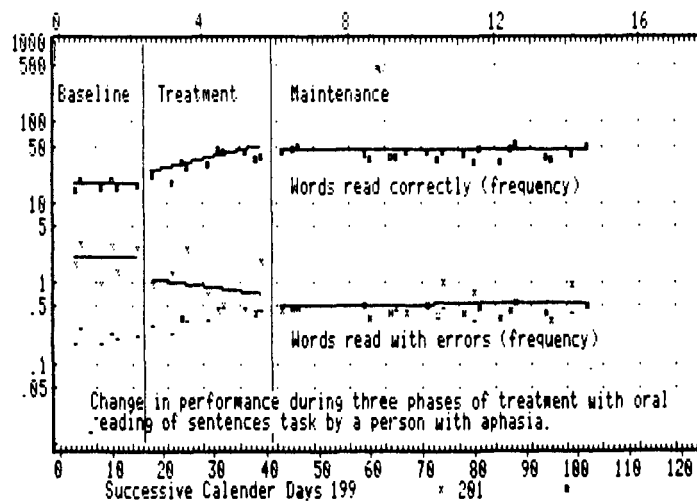


Figure 3. Change in frequency of correct and incorrect word emissions over 4 months of rehabilitation treatment. The dashes indicate the reciprocal of time observed reading each day. This five-cycle semilog display was adapted from the six-cycle Standard Celeration Chart because of the screen-display limitations.

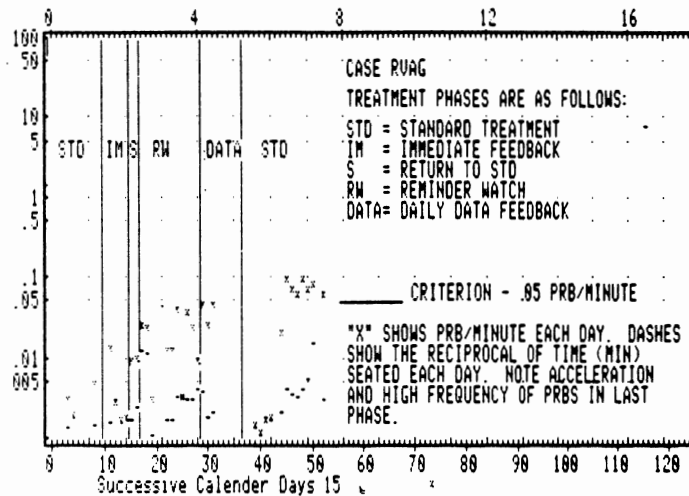


Figure 4. Frequency (count per minute) of pressure-relief behaviors (PRBs) over the course of hospitalization for a person with a spinal-cord injury. The goal is .05 PRBs per minute each day. The dashes indicate the reciprocal of time spent in the wheelchair each day.

has been an issue. In particular, the TLC application has a design goal of continuous, unattended operation for 2 weeks, while monitoring four switches with a temporal resolution of 1 sec and capability of 10K events, powered only by small batteries. Approximately 40 prototypes have been built and tested in pursuit of this goal. While the system is still in development, selected prototypes have been deployed in data collection for research projects. Preliminary reliability data based on experience with 76 patients are available (data from additional patients have yet to be analyzed). In field tests, the TLC system overall has demonstrated a 74.6% data-recovery rate, based on collecting an average of 34 days of data each out of an average of 46 days attempted. Best results were found when resources allowed TLCs to be serviced every 24 or 48 h, as opposed to attempting 1 or 2 weeks of continual data collection. While these results are encouraging, further work on operational reliability is needed.

Accuracy has been addressed on several levels. During development, applications are repeatedly tested by applying known behaviors at specified intervals and verifying system output against the known inputs. When application software yields inappropriate outputs, it returns to the programmer for further attention. Because assembly language is global, a complete test cycle is needed every time any change is made. Testing and debugging is thus a labor- and time-intensive process. However, applications have not been released until all testing is completed with perfect accuracy. Hardware is similarly tested and recycled; the TLC system in particular has benefited from extensive field testing, although two problems persist. First, during long data-collection sessions, persons other than the subject can create a spurious record by sitting in an instrumented wheelchair. Second, the exact time and pressure parameters minimally required to induce a sore or to prevent one are unknown; therefore, we set the

TLC's transducers to record PRBs that reflect a clinical judgment of safety.

Data Files

All data files have a name, a header, a comment, and a data stream. Any legal ProDOS file names can be used. Sorts and searches can be performed on five different fields in the file name, so file names are controlled by MOMS. Each application is coded to the 11th character in the file name. The system uses the file date as kept in ProDOS, and files can also be sorted by date. The header contains a reference time of day, patient and clinician names, and pointers to the end of the comment and the end of the data stream. Applications have a predefined temporal resolution (ranging from 1 to .001 sec) for events, and an identity byte for temporal resolution is also kept in the header. The comment is a text area for clinical notes. The data stream is a series of three byte words in which one byte is devoted to a numeric code for an event and two bytes keep the time of day at which it occurred as an offset from the time in the header. A maximum of 254 classes of events can be recognized; 240 is reserved as a clock rollover to permit unlimited-length data files, and 255 is used as a "wild card" in data analysis. Each application also has a dictionary that contains labels of up to five characters for each event code. During data analysis, the labels are displayed with the numeric codes.

Data Analysis with DADS

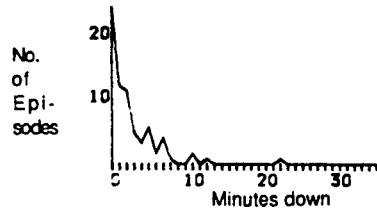
Figure 6 shows the DADS main menu and first-level submenus. Many items are self-explanatory; selected menu functions are discussed below, followed by selected command-mode functions.

The basic data listing is the time of day, the numeric event code, and the alphabetic label for the event. Two data editors are supported; data entries can be modified,

DAILY TLC REPORT

Daily Report: **VAG50** Tech: **WJ**
 Totals: () Ups > () Minutes up () Minutes down () Ups/minute down
 Distribution of Downs: (Number of episodes by minutes down) Worst Push-Up Periods * Best Push-up Periods +
 From Duration Activity From Duration Activity

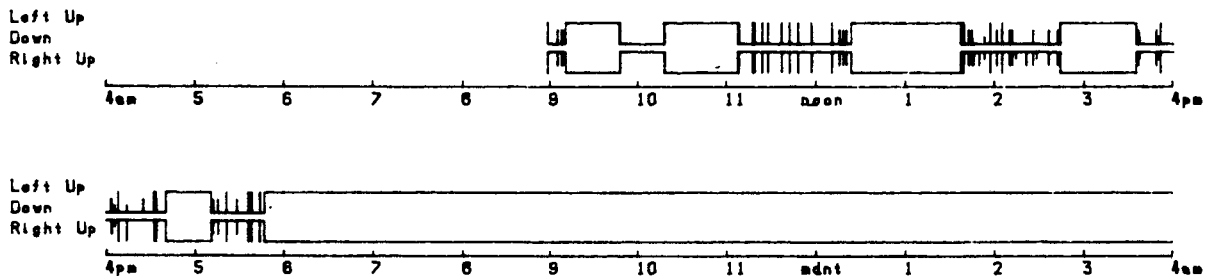
09:54:35 23



Date/File

EVENT LOG

PLOTTED: 17-NOV-88 21:02 VAG50NOV17TH825



Comments:

By: _____ At: _____

Figure 5. Pressure-relief behavior (PRB) over a 24-h period for the person shown in Figure 4. The center tracings of the Daily TLC Report show left and right buttock pressure. This subject began sitting just before 9:00 a.m. (tracings together) and was out of the wheelchair from about 9:10 to 9:45 (tracings apart). The spikes (such as those between 11:10 and 12:20) each indicate a PRB greater than 5 sec in duration. Shorter spikes indicate less pressure relief. Asymmetry indicates pressure relief on one buttock but not the other. The "Distribution of Downs" in the upper left shows the number of episodes of sitting at each indicated duration. Only one episode exceeded our criterion of at least one PRB every 20 min and made the "Worst Push-Up Periods" list; "Best Periods" was not implemented when these data were collected.

deleted, or added to a file, and the file can be saved with a new name. Either across the whole file or within a defined range, data can be counted, frequencies (count per minute) calculated, or elapsed times between events determined. These functions work with either single events or user-defined sequences of events. Events can be grouped for analysis using < > operators on the numeric codes. A "wild card" is supported so that one can search for sequences without specifying all behaviors in the sequence. This feature allows easy lag analysis of data. Finally, analysis of segments of data files bounded by specified events is supported.

DADS also has a "command prompt mode" from which all functions are available, and switching between functions is more rapid. The basic data listing of time of day, numeric code, and alpha label can be expanded to include the latencies between events. The command mode also supports macros. The user can write a macro in Apple-

Works (or a similar word processor) containing a series of DADS commands and print it to disk. Then a batch of files can be selected and the macro called. DADS will step through the files, executing each command. This feature is useful for printing summaries of a series of data files and other repetitive tasks.

Two general forms of graphic data analysis are supported. One graphic provides for cumulative and event records of a single session of data (Figure 2). A "layout" screen allows the user to specify what classes of behavior should be displayed, how they will be plotted, which segments of the file to plot, and so forth. Event and cumulative records can be mixed on the screen as desired. Data are automatically presented at a resolution that fills the screen, but the display can be adjusted from 7 sec to 24 h per screen and data can be scrolled in both directions. A plotter is also supported (Figure 5) for event records of 24-h patient-behavior monitoring.

Map of DADS 6.0 Main Menu (top row) and Submenus (lower rows). DADS is reached from an application Main Menu.

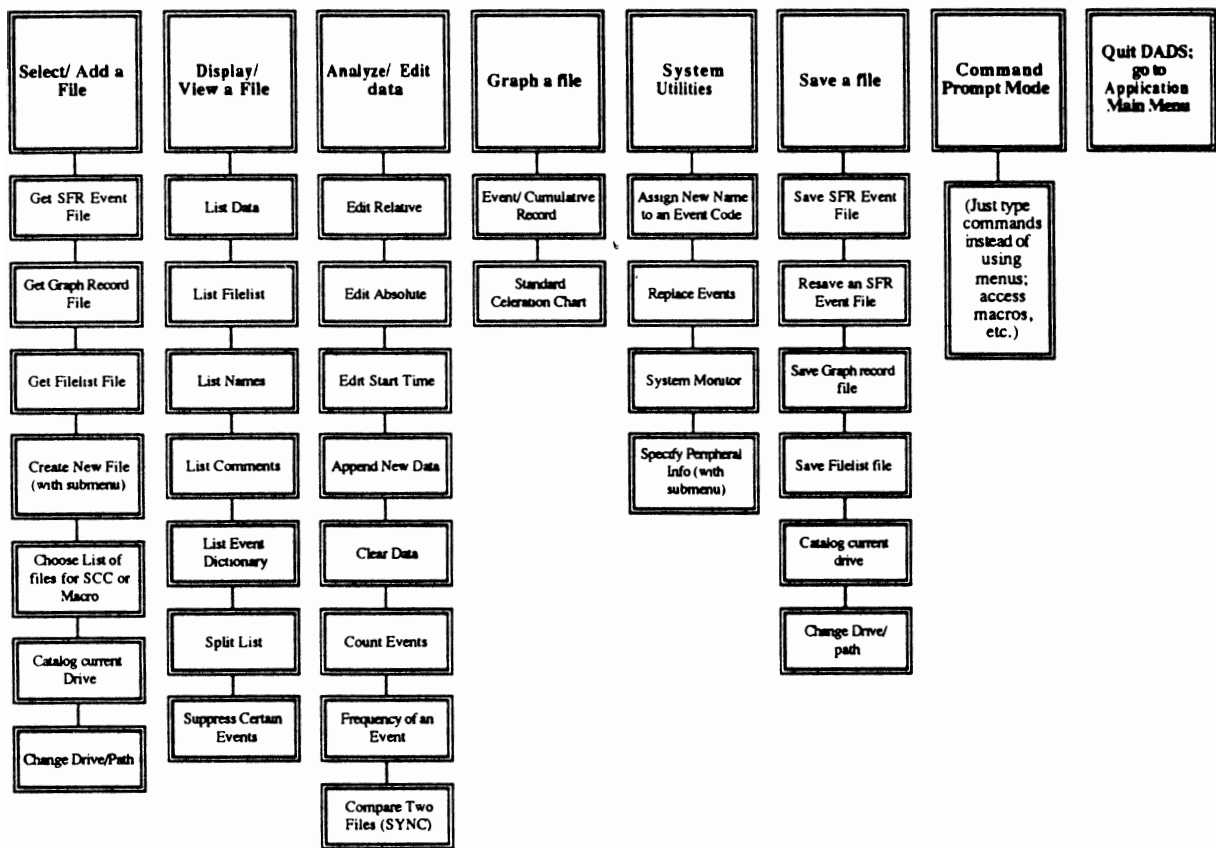


Figure 6. The DADS main menu and first-level submenus.

The other graphic, a modification of the Pennypacker, Koenig, and Lindsley (1972) Standard Celeration Chart (SCC) provides between-sessions data analysis. The user selects which sessions and behavior to plot and may plot data from only selected portions of the sessions or limit the eligibility of data for analysis in several ways. DADS compiles the data into a table by accessing the disk files for each session. DADS then generates the plot (see Figures 3 and 4). Labels, experimental phase lines, lines of best fit, and overplots with other behaviors are supported. The SCC was chosen because of its psychometric properties, including direct display of acceleration and deceleration of behavior, wide applicability, uniform display of variability, rational prediction of behavior, and rapid, efficient communication to the viewer (Johnston & Pennypacker, 1980).

Several printers are supported. Both list-type and graphic analyses are easily printed. With the exception of Fig-

ure 5 (done on a plotter), other data illustrated here were printed on an Apple ImageWriter II.

Plans

As currently implemented, PRS runs on the Apple IIe, IIc, and IIgs family, a platform that is no longer manufactured. After finishing the applications for the Apple platform that are in final stages of development, transferring the system to the IBM PC and Macintosh operating systems is the next step. Improved capability to analyze data collected by other systems and import/export of data to other systems is also planned, following Schneider (1991). Ultimately, commercial release of the PRS modules is desirable so that user resources can support distribution and improvement.

One PRS application, the CAS (distributed by Parrot Software, State College, PA) is now commercially available and has been competitively recognized; it was a

Regional Finalist in the 1991 Johns Hopkins University National Search for Computing to Assist Persons with Disabilities. Thus within this venue it was rated one of the 100 best computer applications in the U.S.

Implications

Behavioral treatment in rehabilitation is a task of great magnitude. The PRS is a beginning that provides a model for absolute measures of a range of behaviors. In addition to its direct benefits to patient outcomes, the PRS has broader implications.

First, because all PRS data collection and analysis uses an absolute time-based system, it can provide clinicians from different disciplines that are responsible for an individual with a ratio-quality data language with which to describe patient behavior and responses to intervention. Wide usage of these modules and this sort of communication may foster better staff coordination of interventions and improve rehabilitation efficiency.

Second, routine usage of the PRS will provide each clinician with an accessible personal database of outcomes of clinical procedures. Such data can be used as a training aid for self-improvement and for training other clinicians and students.

Third, since PRS measures are absolute, they could be used to determine cost-effect relationships for rehabilitation procedures. To the clinician, such data could be used administratively to better align institutional resource allocation with patient progress. At the level of the institution and third-party payer, such data could also be used to construct a "pay-for-outcome" reimbursement system. For the field, such data can provide more efficient and effective interventions by improving the evaluation of interventions; also, since the system is feedback driven and self-correcting, better interventions will evolve as the system is more widely deployed.

Ultimately, the PRS individual measures can be arranged into a curriculum that includes all important human func-

tions and activities. Such a curriculum would be an improved model of human behavior for psychology as well as for rehabilitation.

REFERENCES

- BERNARD, C. (1957). *An introduction to the study of experimental medicine*. New York: Dover. (Original work published 1865)
- BLEIBERG, J., & MERBITZ, C. T. (1983). Learning goals during initial rehabilitation hospitalization. *Archives of Physical Medicine & Rehabilitation*, *64*, 448-450.
- GRIP, J. C., & MERBITZ, C. T. (1986). Wheelchair-based mobile measurement of behavior for pressure sore prevention. *Journal of Computer Methods & Programs in Biomedicine*, *22*, 137-144.
- JOHNSTON, J. M., & PENNYPACKER, H. S. (1980). *Strategies and tactics of human behavioral research*. Hillsdale, NJ: Erlbaum.
- KAPLAN, R. M. (1990). Behavior as the central outcome in health care. *American Psychologist*, *45*, 1211-1220.
- MERBITZ, C. T., CHERNEY, L. R., & MARQUI, H. P. (1991). *The Communication Analysis System Tutorial*. State College, PA: Parrot Software.
- MERBITZ, C. T., CHERNEY, L. R., & MARQUI, H. P. (1992). The Communication Analysis System. In *Proceedings of the 1991 Johns Hopkins University National Search for Computing to Assist Persons with Disabilities*. Los Alamitos, CA: IEEE Press.
- MERBITZ, C. T., GRIP, J. C., HALPER, A. S., MOGIL, S., CHERNEY, L. R., & BELLAIRE, K. (1989). The Communication Analysis System. *Archives of Physical Medicine & Rehabilitation*, *70*, 118-123.
- MERBITZ, C. T., KING, R. B., BLEIBERG, J., & GRIP, J. C. (1985). Wheelchair push-ups: Measuring pressure relief frequency. *Archives of Physical Medicine & Rehabilitation*, *66*, 433-438.
- MERBITZ, C. T., MORRIS, J., & GRIP, J. C. (1989). Ordinal scales and foundations of misinference. *Archives of Physical Medicine & Rehabilitation*, *70*, 308-312.
- PENNYPACKER, H. S., KOENIG, C., & LINDSLEY, O. R. (1972). *Handbook of the standard behavior chart*. Kansas City, MO: Precision Media.
- ROTH, E. J., MERBITZ, C. T., GRIP, J. C., BOGOLUB, M., MROCEK, K., DUGAN, S., & DONADIO, K. (1990). The Timer-Logger-Communicator gait monitor: Recording temporal gait parameters using a portable computerized device. *International Disability Studies*, *12*, 10-16.
- SCHNEIDER, W. (1991). Equipment is cheap, but the field must develop and support common software for psychological research. *Behavior Research, Methods, Instruments, & Computers*, *23*, 114-116.