31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

New Methodology for Identifying Hierarchical Relationships Among Performance Measures: Concepts and Demonstration in Parkinson's Disease

George V. Kondraske, *Fellow IEEE* and R. Malcolm Stewart

Abstract— **Many different tests that address aspects of human performance have been reported. Yet, critical issues remain. The hierarchical organization of tests, the degree of involvement of different human subsystems, and the relationship between measures are often unclear. General Systems Performance Theory provides the basis for a novel analytic method, termed Nonlinear Causal Resource Analysis, to determine task demands (i.e., analyze tasks) and predict performance in complex tasks using only measures of lower level subsystem performance capacities. Recently, we realized new insights and discovery of a new application of these concepts to address the issues noted. A quasi-objective methodology is presented to identify hierarchical relationships among performance measures. The method is applied to seven different performance measures in a study of Parkinson's Disease subjects (n=197) exhibiting a wide range of disease severity. Resource economic interpretations of experimental data using performance theory concepts were used to define relationships between performance measures and to organize them hierarchically. This method is anticipated to have broad utility for identifying relationships between performance measures.**

I. INTRODUCTION

THERE is a plethora of tests that measure some aspect of human performance [1], [2] and are employed in both medical and non-medical situations. In medicine, issues such as early disease detection [3], determination of disease severity [4], and routine monitoring of rehabilitation progress [5] have motivated the use of more objective and quantitative performance tests. Yet, the hierarchical organization of tests, the degree of involvement of different human subsystems, and the relationship between various test measures are often unclear. It is our observation that this has contributed significantly to the incredible number and types of tests proposed. Here, we present and demonstrate a new analytic method for obtaining insights into the issues noted.

General Systems Performance Theory (GSPT) [6], while applicable to any system-task, was motivated by human performance challenges. For insight into fundamental principles, human system complexity was set aside and focus was placed on simple hypothetical systems and tasks. GSPT provides a first principles, conceptual, quantitative, and hierarchical framework for modeling systems, tasks, and their interface using a modeling abstraction that focuses on the notion of "performance".

With GSPT, systems (and subsystems) are modeled as possessing a set of "performance resources" that reflect the unique qualities that characterize "how well" a given system executes its function (e.g., accuracy, speed, strength, endurance, etc.). The nonlinear, threshold-oriented mathematics of resource economics are incorporated, which simply states that resource availability (R_A) must exceed resource demand (R_D) for "success" of a given system in a given task (i.e., $R_A \geq R_D$). Multiple performance resources are drawn upon by tasks; this thus extends to the logical combination of resources; i.e., "sufficiency" (≥) is required for Resource I AND J AND L, etc. The concept of a performance capacity envelope is derived from this representation, the volume of which represents *the capacity of the system to perform tasks that make demands on performance resources that form the multi-dimensional performance space.* A task analysis, performance modeling and performance prediction methodology dubbed Nonlinear Causal Resource Analysis (NCRA) was also developed using GSPT [7]-[9]. NCRA not only estimates the level of performance in a higher level task (HLT) supported by a set of lower level or basic performance resources (BPRs), but also identifies which BPRs limit HLT performance.

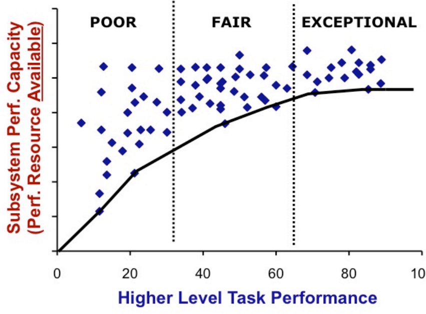

Fig. 1. Typical scatter plot obtained when the measured performance capacity for a given subsystem drawn upon in a higher level task is plotted against performance in the more complex task.

Fig. 1 provides insight into relevant aspects of the GSPT/NCRA methodology. Here, the performance capacity of a human subsystem that is clearly involved in the execution of a more complex HLT is plotted against the level of performance achieved in the HLT. Thus, each point represents the subject's performance capacity at two different hierarchical levels. While it had been hitherto common to fit a line "through the data" (i.e., a correlation

Manuscript received April 23, 2009. This work was supported in part by the Horace C. Cabe Foundation and donations from Mr. Charles Sitter.

George V. Kondraske is with the Human Performance Institute, University of Texas at Arlington, Arlington, TX 76019-0180 USA (phone: 817-272-2335, fax: 817-272-2253, e-mail: kondraske@uta.edu).

R. Malcolm Stewart is with the Human Performance Lab, Presbyterian Hospital of Dallas, Dallas, TX 75231 USA (e-mail: R.MalcolmStewartMD@texashealth.org).

978-1-4244-3296-7/09/$25.00 ©2009 IEEE

approach), GSPT suggests a different interpretation. Specifically, with resource economic thinking, one asks the question: "What is the minimum amount of the basic subsystem capacity *required* for a subject to be classified in the 'exceptional' HLT category"? This logic extends to the "fair" and "poor" categories, resulting in a curve termed the "resource demand function" (i.e., lower boundary of scatter-plot points). The upper-left triangular distribution of data occurs when both measures are defined such that a larger numerical value represents better performance (i.e., defined using a resource construct as per GSPT) [6] and the vertical axis measure represents the availability of a BPR utilized to achieve the HLT performance plotted along the horizontal axis. The validity of these concepts has been demonstrated experimentally, for example in [7]-[9].

In the current study, scatter plots are generated using combinations of two selected performance measures. They are then examined to determine the data distribution pattern and relative hierarchical relationship between the plotted measures. Typical outcomes are illustrated in Figs. 2-4.

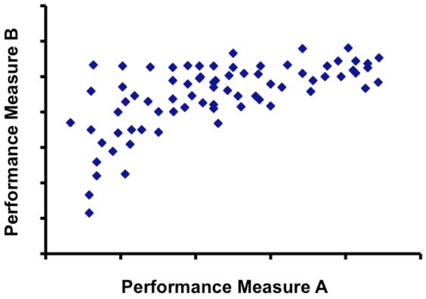

Fig. 2. Performance Measure B is a "more basic" performance resource drawn upon to achieve the performance exhibited in the "more complex" HLT task represented by Performance Measure A.

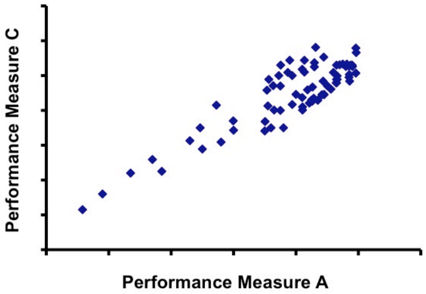

Fig. 3. Performance Measure C is no different than A; i.e., they likely represent different measures of the same performance resource.

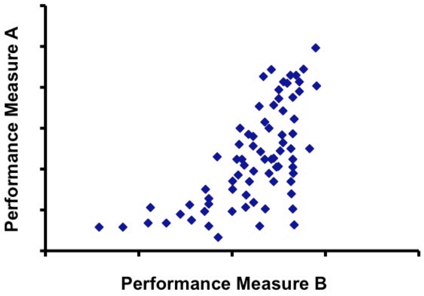

Fig. 4. (The inverse of Fig. 2). Performance Measure B (now on the horizontal axis) is the "more basic" or required resource supporting performance in the "more complex" task (Performance Measure A).

II. METHODOLOGY

A study was conducted using a de-identified data set from our database.

A. Subjects

Data from 197 Parkinson's Disease (PD) subjects (age 34-90 years, 74% male, 26% female), who each had data for a specific set of the measures of interest, was selected.

B. Performance Measures

Two types of performance tests were included. Paper and pencil tests consisted of the Mini Mental Status Examination (MMSE) and a Clock Drawing (CD) task. The other type included selected subsystem performance capacity measures acquired with Models BEP I and BEP II (Human Performance Measurement, Inc., Arlington, TX). The latter [10] were designed to conform to Elemental Resource Model constructs [6]-[8] and thus reflect measures of performance resource availability.

1) MMSE: This test [11] is widely used as a quick and simple means to screen for cognitive loss. It tests orientation, attention, calculation, recall, language and simple motor skills.

2) CD: This test [12]-[14] provides a quick and easily administered evaluation of executive fuction (not completely evaluated by the MMSE), spatial function, and neglect. As implied, subjects are required to draw a clock, putting in all the numbers and setting the hands at ten past eleven. Clock drawing was scored using the Mendez *et al* method [14].

3) Simple Visual-Arm Response Speed (VARS): This is a classic visual reaction time test stressing speed of response to a visual stimulus and renamed to reflect the resource measured. Using the BEP I (Fig. 5), the subject places their fingertips on the central "Home" touch sensor and waits for the visual stimulus (all eight lights simultaneously lighted). The result is the reciprocal of the measured reaction time (responses/s). The average of the best 3 of 5 trials is used as the measure of this capacity.

Fig. 5. The BEP I Central Processing and Upper Extremity Motor Control Performance Capacity Measurement System.

4) Visual Information Processing Speed (VIPS-4C): This is similar to the VARS test, but with higher load on visual information processing as a function of the number of choices presented. A 4-choice load was used. In this test, one of the four center lights (LED3 thru LED6) is randomly selected (equiprobable) to light. The subject must react as quickly as possible by moving their hand from the "home" touch sensor to the touch sensor in front of the lighted LED.

The reciprocal of the reaction time (stimuli/s) is computed. The average of the best 9 of 12 trials (randomly distributed across the four possible stimuli) is used as the final measure.

5) Finger Tapping Speed (FTS): This is an example of a "rapid alternating movement" test, a general type of test used in neurology to screen the integrity of neuromotor subsystems. For FTS, the subject taps any one of six touch sensitive regions of the BEP I instrument (labeled B1-B6 in Fig. 5) "as fast as possible" with his/her index finger for 10s. Subjects are instructed to restrict motion so that it occurs only about the MCP joint (knuckle) of the index finger. Software detects the first tap and tracks the number of taps as well as the elapsed time for key events (e.g., finger contact with or removal from a touch sensor). Various performance measures are calculated, including average tapping speed (taps/s), the primary measure in the present study. The average from two trials is used as the final result.

6) Visual-Spatial Short-Term Memory Capacity (VSSTMC): Eight touch sensors (A1-A8) and corresponding lights (LED1-8) arranged in a semi-circle on the BEP I are used. The number of items correctly recalled is determined after a stimulus consisting of a random sequence of the spatially distributed lights is presented. Responses are communicated manually via touch sensors. Each trial consists of multiple sequences with increasing memory stress (up to 20 items) to determine maximum capacity. The best of two trials is used as the final result.

7) Neuromotor Channel Capacity (NMCC): This test is based on the speed-accuracy tradeoff known as Fitts' Law with additional conceptual support from GSPT. NMCC has been argued to be an objective, conceptually based measure of coordination [15]. The six BEP I touch sensors nearest to the subject (label B1-B6 in Fig. 5) are used. B2 and B5 (i.e., the narrow regions) serve as "targets", while other flanking regions record errors when the subject attempts to hit a target. During a 10s maximal performance reciprocating (right-to-left, left-to-right, etc.) "lateral reach-and-tap" task that stresses speed and accuracy, NMCC (bits/s) is computed. A familiarization trial is provided first. The best of two subsequent trials is used as the final result.

For lower extremities, the BEP II device was employed. The paradigm is analogous to that described for the BEP I. A four-limb NMCC measure (4LNMCC) is derived as the mathematical product of the left and right upper and lower extremity channel capacities [15].

C. Performance Theory-based Graphical Data Analysis

Hierarchical relationships between measures were systematically evaluated. Scatter plots were generated using the following process: 1) a measure was selected and used as horizontal axis data (i.e., making it a candidate representing HLT performance; refer to Fig.1), 2) each of the remaining six measures was then used as vertical axis data in separate scatter plots (i.e., making each a candidate representing a "more basic" subsystem performance capacity), and 3) steps 1 and 2 were repeated for each of remaining measures. This yielded 42 scatter plots, half of which were simply inverses of other plots in the set. The data distribution pattern in each plot was subjectively evaluated (using Figs. 2-4 for reference) to determine the relative hierarchical relationship (i.e., identify which represents a lower level performance resource and which reflects a HLT performance capacity). The collective results were then logically assessed to determine an overall hierarchy among all measures.

III. RESULTS

Two measures (VARS and VIPS-4C) produced scatter plots with a pattern as shown in Fig. 3. All other plots except those involving the FTS measure, exhibited patterns shown in Fig. 2 or 4. Examples are shown in Figs. 6-8.

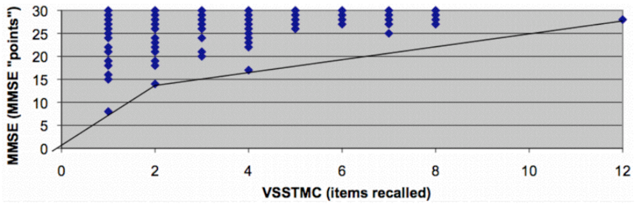

Fig. 6. Pattern suggests that MMSE is a "performance resource" for VSSTMC.

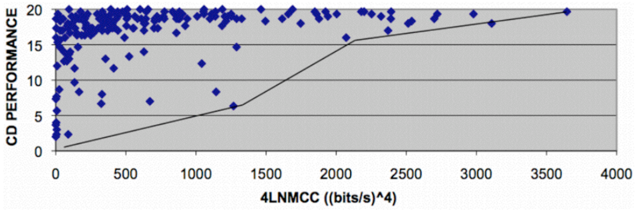

Fig. 7. Pattern suggests that CD performance is a "performance resource" for 4LNMCC.

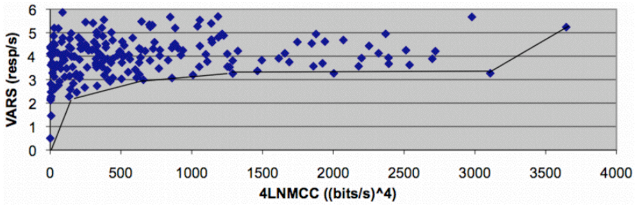

Fig. 8. Pattern suggests that VARS is a "performance resource" for 4LNMCC.

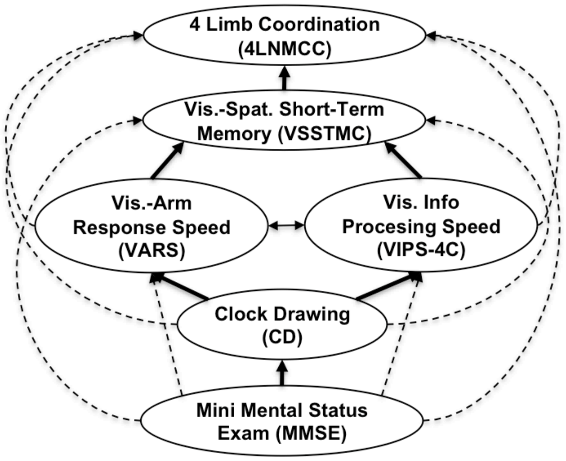

Fig. 9. Illustration of relationships identified using the proposed method. Dark lines/arrows represent the primary hierarchical order identified. Dashed lines/arrows represent situations where the measures from the lower positioned tests were determined to be "basic performance resources" with respect to the measures to which arrows point.

The relationships among all measures (except FTS) are illustrated in Fig. 9. From lowest to highest hierarchical level, the following order was found: MMSE, CD, Visual Information Processing Speed (VIPS-4C) and Visual-Arm Response Speed (VARS) (both found to be at a similar level), Visual Spatial Short-Term Memory Capacity (VSSTMC), and 4-Limb Neuromotor Channel Capacity (4LNMCC).

IV. DISCUSSION

No clear type of relationship was found between Finger Tapping Speed (FTS) and any other measure, possibly implying that it stresses some unique performance resource.

Data points representing a high level of subsystem performance resource availability while HLT performance is poor (e.g., as in Figs. 6-8) are explained by: 1) noting that *multiple* lower level performance capacities are combined to accomplish the HLT and 2) realizing that a capacity other than that in the plot is limiting HLT performance.

The type of plot shown in Fig. 1 is a cornerstone of GSPT and NCRA. Some may question the assignment of independent (HLT performance) and dependent (a lower level performance capacity) variables since HLT performance can be viewed to depend on lower level subsystem capacities. However, during GSPT development HLT performance was viewed as a known quantity (often measured with subjective, Gestalt methods in complex tasks). A key question of interest was, *"How much of a given lower level performance resource is required (or utilized) to obtain a particular level of HLT performance?"* This "unknown quantity" was viewed to be a function of HLT performance; i.e., a greater amount would be required at higher levels of HLT performance (thus the name "resource demand function"). We recently learned of an interesting situation in the field of economics, no less, pertaining to what may also appear to be unconventional labeling of axes [16]. Alfred Marshall popularized what is widely considered the most important graph in economics and the hallmark of supply-demand modeling, that of price versus quantity (presented earlier by F. Jenkin in 1870). While some experts consider Marshall's choice of axes to simply be incorrect [16], others (e.g., [17]) have studied the history and have put forth logical explanations.

The MMSE and CD address cognitve function and are used for this purpose in PD. While objective tests that focus on motor performance are also used, they involve cognition as well. For example, compliance with test instructions for a test such as NMCC requires understanding and memory. Thus, a good result in a test such as 4LNMCC can be used to infer a minimum level of a cognitive performance capacity.

Clearly, *any* task can be the basis for a human performance test. This circumstance, along with others that have been well noted [6], contributes to the lack of a relatively small set of broadly accepted standardized measures for human performance characterization. The human system's architecture is *the* common denominator across all diseases and injuries as well as sport, work, and recreational endeavors. Thus, it is not unreasonable to envision a set of such measures. Methods such as that presented here may help foster better test design, interpretation, and standardization.

V. CONCLUSION

Using GSPT and NCRA concepts, a new method to explore hierarchical relationships among performance measures has been derived and experimentally demonstrated. Subsystem performance capacity tests that do not focus on cognition could have value in its evaluation. The methodology and logic provide a new tool to study human systems and tasks, especially test tasks and the subsystems they target.

REFERENCES

[1] V. J. Gawron, *Human Performance Measures Handbook.* Philadelphia: Lawrence Erlbaum Associates, 2000.

[2] A. R. Potvin, W.W. Tourtellotte, J.H. Potvin, G.V. Kondraske, and K. Syndulko, *The Quantitative Examination of Neurologic Function*, Boca Raton, FL: CRC Press, 1985.

[3] A. M. Johnson, P. A. Vernon, Q. J. Almeida, L.L. Grantier, R. Singarayer and M. S. Jog, "Screening for Parkinson's disease with response time batteries: A pilot study," *BMC Medical Informatics and Decision Making*, vol. 4, pp. 14, 2004.

[4] G. V. Kondraske and R.M. Stewart, "Quantitative characterization of disease severity in diseases with complex symptom profiles," *in CD-ROM Proceedings of the 28th Int. Conf. of the IEEE Eng. In Med. Biol. Soc.*, New York, 2006, pp. 3966-3969.

[5] G. V. Kondraske, "Measurement tools and processes in rehabilitation engineering," in *The Biomedical Engineering Handbook, 3rd ed.: Biomedical Engineering Fundamentals,* J. Bronzino Ed. Boca Raton: CRC Press, Taylor & Francis, 2006, pp. 73.1-73-16.

[6] G. V. Kondraske, "The elemental resource model for human performance," in *The Biomedical Engineering Handbook 3rd Ed.: Biomedical Engineering Fundamentals.* J. Bronzino, Ed. Boca Raton: CRC Press, Taylor & Francis, 2006, pp. 75.1-75.19.

[7] M. T. Gettman, G. V. Kondraske, O. Traxer, K. Ogan, C. Napper, D. B. Jones, M. S. Pearle, and J. Cadeddu, "Assessment of basic human performance resources predicts operative performance of laparoscopic surgery," *J. Amer. Col. Surg.*, vol. 197(3), pp. 489-496, 2003.

[8] C.A. Fischer, G.V. Kondraske, and R.M. Stewart, "Prediction of driving performance using nonlinear causal resource analysis," in *CD-ROM Proceedings of the 24th Int. Conf. IEEE Eng. in Med. Biol. Soc.*, Houston, 2002, pp. 2473-2474.

[9] C. R. Innes, R. D. Jones, J. C. Dalrymple-Alford, S. Hayes, S. Hollobon, J. Severinsen, G. Smith, A. Nicholls, and T. J. Anderson, "Sensory-motor and cognitive tests predict driving ability of persons with brain disorders," *J. Neurol. Sci.,* vol. 260(1-2), pp. 188-98, 2007.

[10] G. V. Kondraske, "A PC-based performance measurement laboratory system," *J. of Clinical Engineering,* vol. 15, no. 6, pp. 467-478, 1990.

[11] M. F. Folstein, S. E. Folstein, and P. R. McHugh, " 'Mini-Mental State' a practical method for grading the cognitive state of patients for the clinician, " *J. Psychiat. Res.*, vol. 12, pp. 189-198, 1975.

[12] T. Sunderland, J. L. Hill, A. M. Mellow e*t al.*, "Clock drawing in Alzheimer's disease: a novel measure of dementia severity," *J. Am. Geriatr. Soc.*, vol. 37, pp. 725-729, 1989.

[13] J. Angela, T. Shirely, and B. Victoria, "The value of clock drawing in identifying executive cognitive dysfunction in people with a normal Mini-Mental State Examination score," *Can. Med. Assoc. J* , vol. 167, pp. 859-8, 2002.

[14] M. F. Mendez, T. Ala, and K.L. Underwood, "Development of scoring criteria for the clock drawing task in Alzheimer's disease," *J. Am. Geriatr. Soc.*, vol. 40, pp. 1095-1099, 1992.

[15] G. V. Kondraske and R. M. Stewart, "Neuromotor channel capacity as an outcome and tracking measure in Parkinson's Disease*," in CD-ROM Proc. of the 24th Int. Conf. IEEE Eng. in Med. Biol. Soc.*, Houston, 2002, pp. 2471-2472.

[16] R. G. Lipsey and K. A. Chrystal, *Economics,* 11th ed. New York: Oxford University Press, 2007, pp. 40.

[17] S. H. Gordon, Scott H. "Why did Marshall transpose the axes?," *Eastern Economic J.*, vol. 8(1), pp. 31-45, 1982.