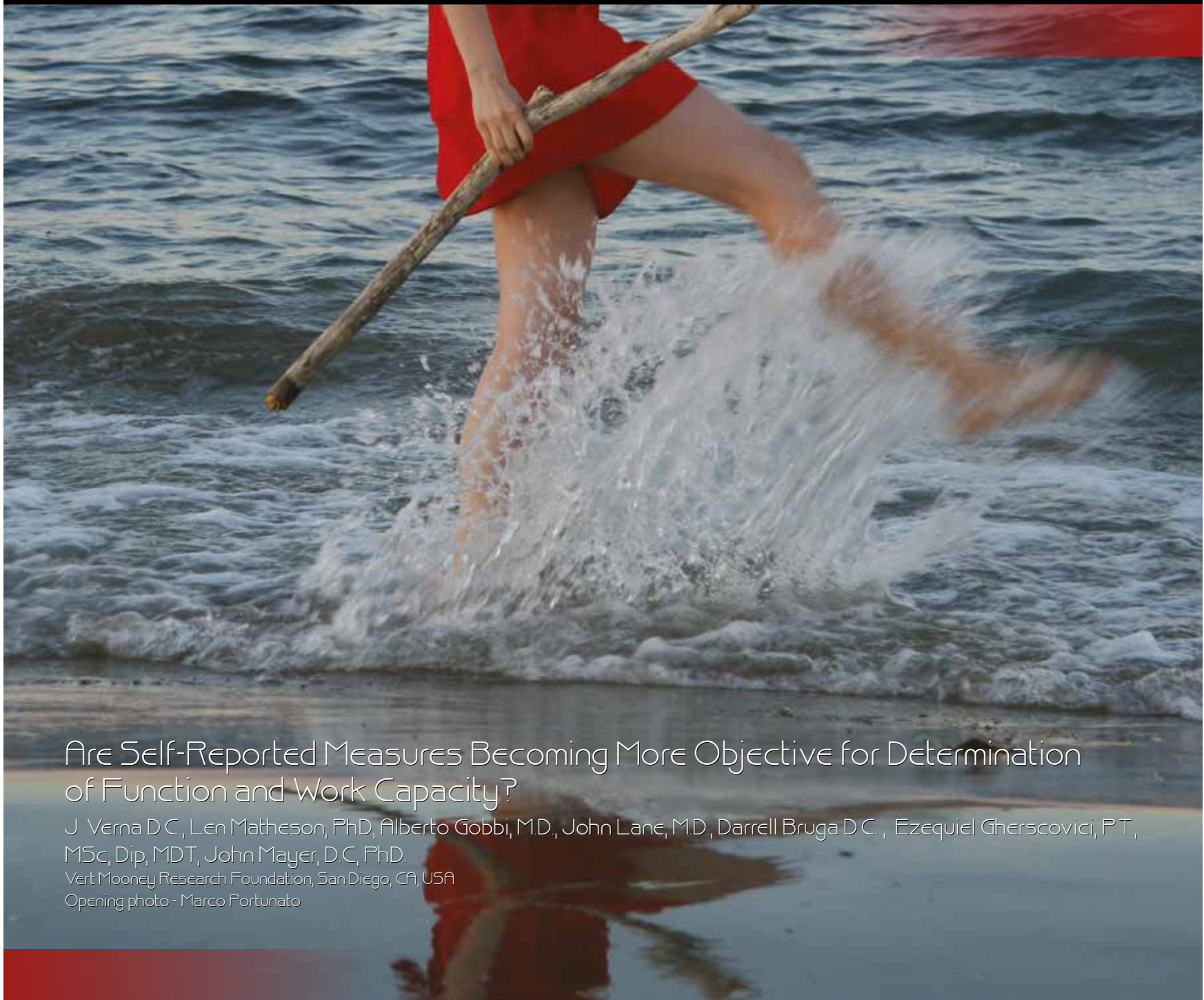


Self-Reported Measures



Are Self-Reported Measures Becoming More Objective for Determination of Function and Work Capacity?

J. Verna D.C., Len Matheson, PhD, Alberto Gobbi, M.D., John Lane, M.D., Darrell Bruga D.C., Ezequiel Ghercovici, P.T., MSc, Dip, MDT, John Mayer, D.C, PhD.
Vert Mooney Research Foundation, San Diego, CA, USA
Opening photo - Marco Fortunato

ABSTRACT

Patient self-reported outcome (PRO) measures are routinely used in healthcare to measure baseline status, progress and treatment outcomes. PRO measures have become integrated into research and clinical practice for physical medicine. Their use is recommended by most professional organizations and stakeholders. However, many of the most popular measures were developed without modern technologies, particularly the mathematical application of item response theory (IRT). The use of these legacy measures in research and the

clinic has become suspect due to the rapid advancement of technology. Psychometric studies of several new self-report instruments that use IRT blur the line between subjective and objective measurements. This IRT technology provides useful information regarding a patient's function and work capacity safely and efficiently. The purpose of this article is to summarize how the Multidimensional Task Ability Profile (MTAP) was developed using IRT and integrated into physical medicine rehabilitation practice.

INTRODUCTION

The importance of utilizing standardized instruments measuring health and functional status cannot be overstated. Utilization of patient reported outcome measures (PROs) has been advocated globally by physical medicine providers for many years^{1,2}. The National Institutes of Health has funded the Patient Reported Outcome Measurement Systems (PROMIS) in 2004 and re-funded it in 2010³. PROMIS has developed more than 40 instruments using item-response theory, which provides advantages in quantification of function over classical test theory.

In contrast to invasive physical tests, the routine use of PROs in clinical practice provides a mechanism safely, efficiently, and inexpensively. PROs provide large amounts of information about patient function, health status, and work capacity⁴⁻⁸.

Advocates of self-reported measures suggest that PROs enhance communication with patients and help direct plan of care and treatment algorithms^{4,8}. On the other hand, some suggest that such measures possess inherent deficiencies regarding a lack of accuracy due to subjectivity⁹⁻¹¹(1,6,7,8). Further, practical issues have been cited as common reasons for the lack of implementation of outcome measures into routine clinical practice include: extensive time length for patients to complete testing, length of time for clinicians to analyze and interpret data, and inability of patients to be tested independently due to reading or language problems⁵.

Traditionally, most self-reported measures in physical rehabilitation were developed utilizing the classical test theory without item calibration or proportional mathematical comparisons of total scores. In the effort to enhance the accuracy of self-reported measures, advances in technology have recently been adopted. Development of testing methods that originated in the fields of education and psychology provide item calibration, allowing proportional evaluation of total scores.

ITEM RESPONSE THEORY AND RASCH ANALYSIS

A major advancement in PROs include the incorporation of Rasch analysis, which is a method to calibrate items and utilizes Item Response Theory (IRT), which improves the psychometric reliability and validity for these measures^{9,12}.

Although item calibration and rating scale calibration are widespread in the field of education, the need for such calibration has only recently been appreciated in healthcare⁹. Rasch analysis includes the ability to predict how a subject or evaluatee would likely answer or respond to certain items to a high degree of probability^{13,14}. The Rasch IRT provides an “Infit” score and an indicated response different from expected response patterns on items near the ability level of the evaluatee. This Infit score provides a method to examine the reliability of the match of the evaluatee to the items within range of the evaluatee’s true ability. The “Outfit” score is sensitive to items that are outliers, either very easy or very difficult, compared to the evaluatee’s ability score. The Outfit score reflects unusual responses that are at the extremes of the evaluatee’s ability score, likely indicating misunderstanding of the item. These procedures allow for the proportional calibration of ordinal self-report items improving reliability and validity of the instruments and allow higher levels of sensitivity and specificity^{9,15,16}. Various methods of Rasch and other response theories are currently applied in healthcare to improve the psychometric reliability and validity of measures and are being used, for example, by the U.S. National Institutes of Health (NIH) Patient-Reported Outcomes Measurement Information System (PROMIS project)¹⁶⁻¹⁸. Additionally, IRT calibration methodology can be used to cross reference and integrate self-reported data with objective physical performance testing^{12,19}. The successful incorporation of IRT,

Rasch analysis and the integration or cross validation to actual objective physical performance measures leads to a more stable and objective self-reported outcome tool. Novel healthcare self-reported measures that incorporate IRT, Rasch analysis and computer adaptive technologies in an effort to enhance validity and reliability are being recognized and their routine use is advocated^{3,16,17}.

SUBJECTIVE VERSUS OBJECTIVE MEASUREMENTS

Patient-reported outcome measures include subjective items, and are typically considered less credible than measures of ability based on physical performance. The latter are accepted readily because they are considered to be more objective. However, juxtaposition of subjective measurements against objective measures is artificial; these exist along a continuum that is bi-dimensional. Combining both objective and subjective components such as those described in Figure 1 (figure 1). In this model, measures of ability that are based on performance tend to be more objective than subjective, whereas measures of ability that are based on self-report tend to be more subjective than objective, but each type of measure possesses subjective and objective components. For example, functional capacity evaluations (FCEs) and cardiac stress testing are based on volitional performance, which contributes a significant subjective component. On the other hand, self-reported ability measures are based on experience, which contributes a significant objective component⁴. (see Fig. 1)

MEASURING PHYSICAL FUNCTION AND ACTIVITIES OF DAILY LIVING (ADLs)

The Multidimensional Task Ability Profile (MTAP)²⁰ is a web-based and computer-administered patient reported outcome measure designed to assess physical function. The MTAP identifies specific functional limitations and the general effect of these limitations on a person’s ability to work, provide self-care in activities of daily living, and participate in other home or community activities. MTAP uses pictorial activity items with simple text descriptions of common activities of daily living and work activities. The combination of a picture and text allows for more rapid cognitive processing at a lower level of ambiguity than language captions alone. This decreases the evaluatee’s response effort, standardizes the items across patients, decreases errors and is an efficient manner to gather large amounts of data^{9,19}. Moreover, the pictorial activities help to accommodate patients with low literacy levels and assist with cross-cultural adaptation. The use of text alone may lead to language, cultural, literacy issues and limits the use and accuracy of a questionnaire. Precursors to this pictorial activity test questionnaire innovation included the Spinal Function

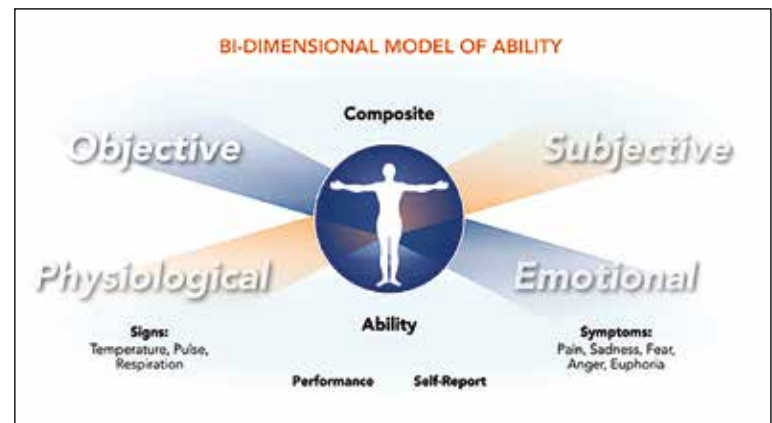
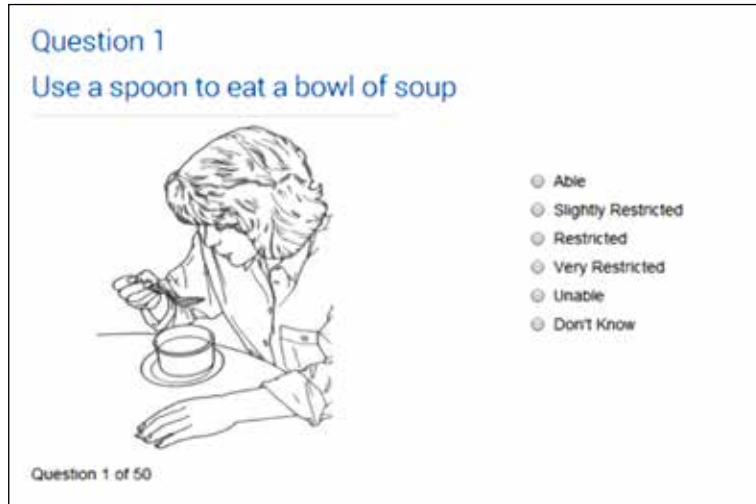


Fig. 1 - Bi-dimensional model of ability

Fig 2 - Example MTAP Instruction Page

You will be asked a series of questions about your current ability. Each question will have a drawing of a task, a short written description, and a rating scale like this:



Look at the drawing and read the written description. Mark your currently ability to perform the task in the written description from “Able” to “Unable”. You do not have to do the task exactly as in the drawing. The drawing is only meant to illustrate the written description. Remember, you do not have to do the task exactly as in the picture. If you have not performed the exact task in the picture, please estimate your ability to perform the task. Work quickly. Do not spend too much time on any one drawing. Your first impression is usually the best.

Sort^{1, 21} and the Hand Function Sort^{22, 23}.

The MTAP incorporates IRT, Rasch analysis, and performance-integrated technology to quantify responses and compare to external performance measures. Simultaneously, mathematical interpretation of the internal consistency of item responses is calculated

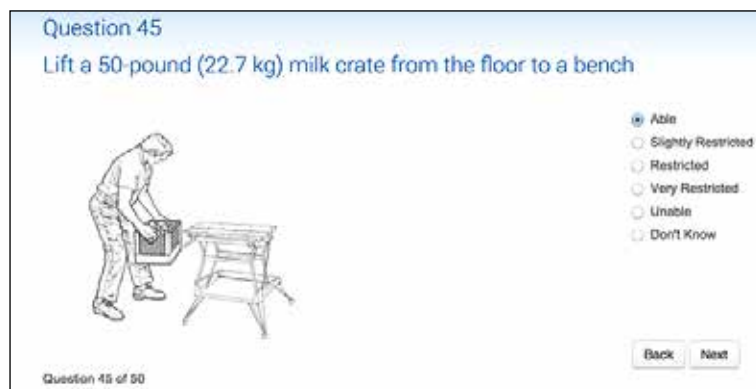


Fig 3 - Example MTAP Question

and scores are benchmarked to common ADLs, work demands and functional capacity evaluations (FCEs)^{19, 20}. (see Fig. 2, see Fig. 3)

Determination of Physical Function and ADLs

The MTAP Report Card results compare the person’s scores to ADLs. The scores will document a functional and ADL category baseline, progress or decline. The report card breaks down common ADLs with categories including Self-Care (very low level) to Heavy Housekeeping, Lawn and Gardening (very high levels). Clinicians can identify deficits in physical function by clinical comparisons and incorporate specific tasks into the treatment algorithms that will enable proper level of function. Secondly, the report card will help communication between the patients and providers and will establish specific goals and a common outcome.

The functional information regarding self-care, cooking and housekeeping ADLs, and category scores helps clinicians make decisions about home health interventions and independent living. For example, someone who cannot perform the majority of self-care will likely need assistance from a caregiver or be transferred to an assisted living facility.

As part of analyzing and interpreting MTAP reports, it is important to determine reliability of the test by interpreting the consistency score, which is listed on the MTAP’s Health and Behavioral Assessment Report. In the absence of polytrauma, INFIT and OUTFIT scores that are in excess of 1.50 indicate unacceptable consistency and require clinical confirmation. Once the consistency of the test is verified, the MTAP results can be accepted into treatment algorithms.

Some reasons for elevated INFIT or OUTFIT scores may be due to the manifestation of adverse psychosocial behaviors. Other possible reasons for inconsistent INFIT and OUTFIT scores may include but not limited to: poor language proficiency, the misunderstanding of items or questions due to poor literacy, cognition difficulties, mismatch of the instrument’s items to the evaluatee’s experiences, or distraction due to commotion in the clinic. Clinical correlation and/or additional testing is advised with high INFIT or OUTFIT scores. (see Fig. 4)

The MTAP analysis is presented in a report used by clinicians to explain test results to the evaluatee. This report provides the current baseline work, along the range from Unemployable, Sedentary, Light, Medium, Heavy and Very Heavy. These are the physical demand characteristics of work (PDC) categories described by the United States Department of Labor^{24, 25}. MTAP scores to be linked to all jobs that are classified according to PDC level. This information can help guide clinical decisions and provides a simple tool to establish work restrictions. When serial testing is performed, work progress can be verified and the work restriction adjusted until a plateau is established. The report also notes the patient’s occupation, job demands, and present work capacity level, and compares the present work ability to the job requirements. This help promotes discussions between the patient and provider regarding functional improvement and return to work. (see Fig. 5, see Fig. 6)

THE MTAP WAS CROSS VALIDATED AND COMPARED WITH “OBJECTIVE” FUNCTIONAL CAPACITY TESTING (FCE)

The MTAP is often used with functional capacity evaluations (FCE). FCE is a method that uses a comprehensive battery of objective performance based tests to determine ability to work and perform activities of daily living. FCEs can help determine decisions about treatment effects, return-to-work and job-placement,

ADL categories

ADL SELF CARE 1.0-2.5 METS	HEAVY HOUSEKEEPING / LIGHT HOME MAINTENANCE 3.5-5.0 METS
bathing or showering, sitting	bathing dog, large
dressing & undressing; standing or sitting	cleaning, house or cabin, general
getting ready for bed, in general	mopping floors
grooming (washing, brushing teeth)	mowing lawn, riding mower
having hair cut or shampooed by someone else	packing/unpacking boxes
low demand sexual activity	picking fruit off trees, picking fruits/vegetables
placing food on plate, cutting food, eating	planting seedlings, shrubs
sitting on toilet, cleaning self	playing active sports with child(ren)
opening containers & taking medication	raking leaves off lawn
talking and eating	trimming shrubs or bushes by hand
LIGHT HOUSEKEEPING 1.5-4.0 METS	HEAVY HOME MAINTENANCE 4.5-6.0 METS
bathing dog, small carpentry, outside	carpentry, outside
child care, seated (dressing, bathing, feeding)	carpentry, refinishing cabinets or furniture
cooking or food preparation	cleaning gutters
gathering clothes to pack, packing suitcase	clearing land, hauling branches
ironing clothes	digging, spading, filling garden, composting
laundry, fold or hang clothes	gardening with heavy power tools
making bed	gardening, general
packing/unpacking boxes, light	hanging storm windows
playing low demand sports with child(ren)	mowing lawn, general
putting away groceries, carrying packages	mowing lawn, walk, hand mower
serving food, setting table	mowing lawn, walk, power mower
knitting, sewing, or wrapping presents	painting, outside home
sweeping floor or sidewalk	painting, papering, plastering, scraping
vacuuming carpet	planting trees
washing dishes	trimming trees
watering lawn or garden, standing or walking	washing fence, painting fence
watering plants	weeding, cultivating garden

MTAP Report Card

Test Physical Therapy **Multidimensional Task Ability Profile** **Junior Hernandez**
REPORT CARD November 10, 2014

Overall Physical Ability

Your current Physical Ability Score is 179 on a 0-200 scale. This demonstrates an improvement of 43% in physical functioning since August 31, 2014.

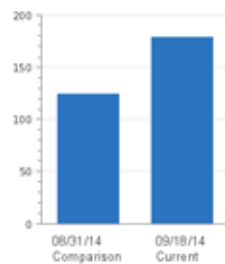
Activities of Daily Living (ADLs)

Your ability to perform ADLs has improved 41% since August 31, 2014.

Ability to Perform

ADL Category	Comparison 08/31/14	Current 09/18/14
Self Care	Many 79%	Almost all 90%
Cooking, Light House Keeping	Almost all 81%	Almost all 90%
Heavy Housekeeping, Light Gardening, Home Maintenance	Many 64%	Almost all 93%
Outside Home Repair, Lawn and Garden Maintenance	Few 27%	Almost all 81%

Progress Between Tests



Physical Demand Characteristics (PDC) of Work

You are able to meet the physical demands for jobs in the Medium work category according to the PDC levels defined by the U.S. Department of Labor. This is an improvement from your PDC level of Light on August 31, 2014.

Improvement Potential

You indicated that you have some restrictions with tasks such as those shown below. Let us know if we do not seem to be adequately addressing problems such as these, or if you have recently experienced difficulty in these areas. Most importantly, let us know if you are experiencing difficulty with other tasks that you regularly perform at work or home. We want to do everything we can to help you improve your physical abilities.



Hammer a large nail into a piece of lumber.



Lift 100-pound (45.4-kg) milk crate from the floor to a bench.

Please let us know how we can continue to assist you. Have a great week!

744 8th Avenue • San Diego, CA • 92101 • (619) 315-5746

Provider Signature: _____

Fig 4 - MTAP "Patient Report Card" and Corresponding ADLs and Typical Energy requirements (METS) in each ADL category.



Multidimensional Task Ability Profile Workability Report

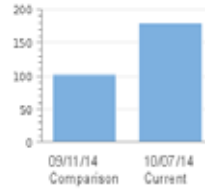
Darrell Bruga
January 31, 2015

Job Title and Work Demands

Your overall Physical Ability score is 179 on a scale of 0-200. This independent test demonstrates an improvement of 74% in physical functioning since September 2014.

Your current job title, Carpenter, requires physical demands in the **Heavy (50-100 lbs.)** work category according to the Physical Demands Characteristics (PDC) levels defined by the U.S. Department of Labor.

Overall Physical Ability



Physical Demand Characteristics of Work

¹ Current PDC Level. ² Target PDC Level

Physical Demand Level	Occasional 0-33% of the workday	Frequent 34-66% of the workday	Constant 67-100% of the workday	Typical Energy Required
Sedentary	10 lbs.	Negligible	Negligible	1.5-2.1 METS
Light	20 lbs.	10 lbs.	Negligible	2.2-3.5 METS
Medium ¹	20 to 50 lbs.	10 to 25 lbs.	10 lbs.	3.6-6.3 METS
Heavy ²	50 to 100 lbs.	25 to 50 lbs.	10 to 20 lbs.	6.4-7.5 METS
Very Heavy	Over 100 lbs.	Over 50 lbs.	Over 20 lbs.	Over 7.5 METS

Workability

Based on today's MTAP testing you are able to meet the physical demands for jobs in the **Medium (20-50 lbs.)** PDC work category. Therefore you are below your occupational demands. The **Medium PDC** level is an improvement of 74% from September 2014. One of the primary rehabilitation goals will be to enable you to safely and dependably return to work or accommodate to modified or full duty activities. A home exercise plan to achieve your functional goals will be included.

Improvement Potential

You indicated that you have some restrictions with tasks such as those shown below. Let your provider know if these problems are not being adequately resolved, or if you have recently experienced difficulty with other tasks that you regularly perform at your work or home.



Lift 100-pound (45.4-kg) milk crate from the floor to eye-level.



Push a full wheelbarrow up a ramp.

Please let us know how we can continue to assist you. Have a great week!

3840 Watt Ave • Sacramento, CA • 95821 • (800) 994-3220
www.lifeteamhealth.com

Provider Signature: _____

Physical Demand Characteristics (PDC) of Work (US Department of Labor 1972).

Physical Demand Level	Occasional 0-33% of the workday	Frequent 34%-66% of the workday	Constant 67%-100% of the workday	Typical Energy Required
Sedentary	10 lbs.	Negligible	Negligible	1.5 - 2.1 METS
Light	20 lbs.	10 lbs.	Negligible	2.2 - 3.5 METS
Medium	20 to 50 lbs.	10 to 25 lbs.	10 lbs.	3.6 - 6.3 METS
Heavy	50 to 100 lbs.	25 to 50 lbs.	10 to 20 lbs.	6.4 - 7.5 METS
Very Heavy	Over 100 lbs.	Over 50 lbs.	Over 20 lbs.	Over 7.5 METS

Fig 5

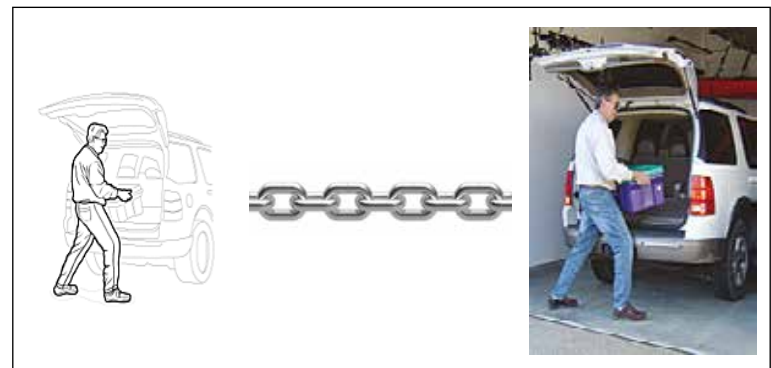


Fig 6 - Pictures allow for calibration and MTAP items are linked to demonstrable physical ability



Note: The subject wears a heart monitor during the FCE to continuously record performance data while they lift, carry and perform various work tasks with blinded weights. The EPIC/ELC possesses published normative performance data that allow comparison within age and gender categories²⁶.

Fig. 7 - EPIC Lift Capacity/ELC

impact on work performance of non-work-related illness and injuries, disability and impairment reporting, treatment plans, and case management²⁶.

The MTAP use of statistical analyses based on IRT and Rasch analysis allows the score to be compared with physical performance objective data obtained during FCE. The MTAP is reliable ($r = 0.98$, $p < 0.05$) and correlates highly with actual physical function as assessed during objective FCE lifting tasks on the EPIC Lift Capacity Test ($r = 0.89$, $p < 0.05$)^{9, 12, 19}. The successful incorporation of IRT, Rasch analysis, and the integration or cross validation to actual objective physical performance measures such as an FCE, leads to a more robust and objective self-reported outcome tool. Additional external linkages cross-referenced include MTAP responses to levels of activities of daily living (ADL), and to instrumental activities of daily living (IADL). (see Fig. 7)

WHAT OTHER SELF-REPORTED MEASURES CALIBRATE ITEMS AND ADDRESS WORK CAPACITY?

In contrast to the MTAP, most popular musculoskeletal self-report measures, such as the 36-Item Short Form Survey Instrument (SF36), Oswestry Pain Disability Questionnaire, and Roland-Morris Disability Questionnaire do not focus on work activities because many patients who use these instruments are not workers. The Disabilities of the Arm, Shoulder and Hand (DASH) questionnaire^{27, 28}, and the QuickDASH questionnaire^{29, 30} include a few work demand items. Thus very little useful information is obtained with respect to the individual's ability to work.

Recently, the US Social Security Administration (SSA) began developing self-report methods to acquire information regarding symptoms and work-relevant physical function as it relates to disability³¹⁻³³. The SSA is utilizing computer adaptive testing (CAT), which employs a computer algorithm that tailors questions to the specific patient functional level from a large pool of items (from low to high levels of function). IRT and Rasch analysis is utilized to calibrate each item in the item pool and to calculate the sequence and administration of the questions. The successful feasibility and development of the SSA self-reported physical function instrument will provide reliable information regarding function and work disability in the near future.

CONCLUSION

Patient reported outcome measures of physical function are becoming standard tools in rehabilitation worldwide. Recently there has been a trend to incorporate new technology in an effort to create more reliable, valid and consistent measures. One such innovation that is rapidly advancing is the incorporation of Rasch analysis and IRT. The use of these advanced mathematical algorithms allows for the calibration of items methodology to greatly improve psychometric properties of these instruments. In addition, this modern technological advancement in testing, allows for the comparisons to actual objective physical performance testing for validation to objective measures.

Despite all of the advancements in testing methodologies, self-reported measures tend to be underutilized clinically, due in part to the misconceptions of subjectivity and difficulty of incorporating into routine clinical practice. However, as measures are refined, the ability for widespread practical application to gather patient functional data safely, efficiently, and inexpensively will expand their future use and popularity.

The MTAP is an example of a patient-reported outcome measure that combines all of the modern technological features with simple drawings of people performing common tasks. This combination of techniques illustrates the ultimate in development and clinical use of self-reported measures. Future research and development will increase the number and diversity of similar measures to help a wide variety of patient populations in various settings. ■

ACKNOWLEDGEMENTS

The authors wish to thank Sharon Gables, ATC for assistance with editing for this study.

J.L. Verna, L.N. Matheson - Vert Mooney Research Foundation, 3670 Convoy St, Suite 101, San Diego, CA 92111, USA
Email: drjverna@aol.com

J.L. Verna - Spine and Sport, Inc, San Diego, CA, USA
L.N. Matheson - EpicRehab, St. Louis, MO, USA
Email: lmatheson@epicrehab.com

A. Gobbi, MD - OASI Bioresearch Foundation, Via G.A. Amadeo 24 20133 Milan, Italy
Email: gobbi@cartilagedoctor.it

J. Lane, MD - 3750 Convoy St. #116, San Diego, CA 92111
Email: John@johnlanemd.com

D. Bruga - LifeTEAM Health, Encinitas, CA, USA
Email: dbruga@lifeteamhealth.com

E.D. Gherscovici, PT, MSc, Dip. MDT - World Spine Care Santa Ana, CA
Email: ezequielgherscovici@g.harvard.edu

J.M. Mayer, University of South Florida, Tampa, FL, USA
Email: jmayer003@tampabay.rr.com

REFERENCES

1. Matheson, L., M. Matheson, and J. Grant, *Development of a measure of perceived functional ability*. Journal of Occupational Rehabilitation, 1993. 3(1): p. 15-30.
2. Gibson, L. and J. Strong, *The reliability and validity of a measure of perceived functional capacity for work in chronic back pain*. Journal of Occupational Rehabilitation, 1996. 6(3): p. 159-175.
3. Cella, D., et al., *The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years*. Medical Care, 2007. 45(5 Suppl 1): p. S3-S11.
4. Mooney, V., et al., *Performance-Integrated Self-Report Measurement of Physical Ability*. The Spine Journal, 2010. 10(1): p. 433-440.
5. Stevens, J.A. and A.J. Beurskens, *Implementation of measurement instruments in physical therapist practice: development of a tailored strategy*.

- Physical therapy, 2010. **90**(6): p. 953-961.
6. Kramer, A. and D. Holthaus, *Uniform patient assessment for post-acute care*. Final report. Aurora, CO: Division of Health Care Policy and Research, University of Colorado at Denver and Health Sciences Center, 2006.
 7. Khorsan, R., et al., *Measures in chiropractic research: choosing patient-based outcome assessments*. Journal of manipulative and physiological therapeutics, 2008. **31**(5): p. 355-375.
 8. Verna, J.L., et al., *Development and reliability testing of Spanish language and English language versions of the multidimensional task ability profile*. Journal of occupational rehabilitation, 2013. **23**(2): p. 220-227.
 9. Matheson, L., et al., *A method to provide a more efficient and reliable measure of self-report physical work capacity for patients with spinal pain*. Journal of Occupational Rehabilitation, 2008.
 10. Michel, A., T. Kohlmann, and H. Raspe, *The association between clinical findings on physical examination and self-reported severity in back pain: results of a population-based study*. Spine, 1997. **22**(3): p. 296-303.
 11. Marras, W.S., et al., *Impairment magnification during dynamic trunk motions*. Spine, 2000. **25**(5): p. 587-595.
 12. Matheson, L., et al. *Calibration of the difficulty of work activities and activities of daily living for patients with spinal pain*. in *North American Spine Society 22nd Annual Meeting*. 2007. Austin, TX.
 13. Bond, T. and C. Fox, *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. 2nd ed. 2007, Mahwah, New Jersey: Lawrence Erlbaum Associates.
 14. Veloza, C.A., et al., *Measurement qualities of a self-report and therapist-scored functional capacity instrument based on the Dictionary of Occupational Titles*. Journal of Occupational Rehabilitation, 2006. **16**(1): p. 109-22.
 15. Reeve, B.B., et al., *Psychometric evaluation and calibration of health-related quality of life item banks: plans for the Patient-Reported Outcomes Measurement Information System (PROMIS)*. Medical care, 2007. **45**(5): p. S22-S31.
 16. Rose, M., et al., *Evaluation of a preliminary physical function item bank supported the expected advantages of the Patient-Reported Outcomes Measurement Information System (PROMIS)*. Journal of clinical epidemiology, 2008. **61**(1): p. 17-33.
 17. Cella, D., et al., *The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005-2008*. Journal of clinical epidemiology, 2010. **63**(11): p. 1179-1194.
 18. Reeve, B.B., et al., *Psychometric evaluation and calibration of health-related quality of life item banks: plans for the Patient-Reported Outcomes Measurement Information System (PROMIS)*. Medical Care, 2007. **45**(5 Suppl 1): p. S22-31.
 19. Mayer, J., et al., *The reliability and validity of a new computerized pictorial activity and task sort*. Journal of Occupational Rehabilitation, 2005. **15**(2): p. 185-195.
 20. Matheson, L., et al., *Multidimensional Task Ability Profile*, 2003, Vert Mooney Research Foundation: San Diego.
 21. Matheson, L. and M. Matheson, *Spinal Function Sort*. 1989, Wildwood, MO: Employment Potential Improvement Corporation.
 22. Matheson, L., M. Matheson, and J. Grant, *Hand Function Sort*, 1995, Employment Potential Improvement Corporation: Wildwood, MO.
 23. Matheson, L.N., V.K. Kaskutas, and D. Mada, *Development and construct validation of the Hand Function Sort*. Journal of Occupational Rehabilitation., 2001. **11**(2): p. 75-86.
 24. U.S. Department of Labor, *The revised handbook for analyzing jobs*. 1991, Washington, DC: Author.
 25. U.S. Department of Labor, *Dictionary of Occupational Titles*. 4th ed. Vol. 1. 1991, Washington, DC: United States Government Printing Office.
 26. Matheson LN, Verna J, Dreisinger TE, Leggett S, Mayer J. Age and gender normative data for lift capacity. *Work* 2014; **49**(2): 257-69.
 27. Hudak, P.L., P.C. Amadio, and C. Bombardier, *Development of an upper extremity outcome measure: the DASH. The Upper Extremity Collaborative Group (UECG)*. American Journal of Industrial Medicine, 1996. **29**(6): p. 602-8.
 28. Beaton, D.E., et al., *The DASH (Disabilities of the Arm, Shoulder and Hand) outcome measure: what do we know about it now?* British Journal of Hand Therapy, 2001. **6**(4): p. 109-18.
 29. Beaton, D.E., et al., *Development of the QuickDASH: comparison of three item-reduction approaches*. Journal of Bone & Joint Surgery - American Volume, 2005. **87**(5): p. 1038-46.
 30. Matheson, L.N., et al., *Reliability of a visual analog version of the Quick-DASH*. J Bone Joint Surg Am, 2006. **88**(8): p. 1782-7.
 31. Marfeo, E.E., et al., *Conceptual foundation for measures of physical function and behavioral health function for social security work disability evaluation*. Archives of physical medicine and rehabilitation, 2013. **94**(9): p. 1645-1652. e2.
 32. Meterko, M., et al., *The Work Disability Functional Assessment Battery (WD-FAB): Feasibility and Psychometric Properties*. Archives of physical medicine and rehabilitation, 2014.
 33. McDonough, C.M., et al., *Development of a self-report physical function instrument for disability assessment: item pool construction and factor analysis*. Archives of physical medicine and rehabilitation, 2013. **94**(9): p. 1653-1660.

