Development of the Variable Dexterity Test: construction, reliability and validity

GONZALEZ, Victor, ROWSON, Jenn and YOXALL, Alaster
<http://orcid.org/0000-0002-0954-2725>

ABSTRACT

INTRODUCTION

This work introduces a dexterity test designed to assess individual types of dexterity utilised whilst carrying out activities of daily living (ADL). Validity and reliability studies for this new test were carried out and the results are shown in this article.

METHOD

Reliability and validity estimates were obtained from 24 healthy participants.
Test-retest and Inter-rater reliability were assessed via ANOVA.
The validity of the test was estimated correlating scores from the VDT with the participant's proficiency to complete each of 4 ADL as well as a gold-standard dexterity test

RESULTS/FINDINGS

The test produces consistent results among a pilot group with both a single assessor (test-retest reliability) and multiple assessors (inter-rater reliability).
Correlations between VDT scores and proficiency to perform ADL were found to be high for most of the subtests.
Correlation between the scores from the Purdue Pegboard Test and the VDT was shown to be high.

CONCLUSION

The VDT proved to be a flexible, reliable, valid tool that approaches the problem of assessing dexterity focusing on activities of daily living for the pilot group. Validity and reliability estimates show encouraging values, proving that the VDT can be used as an accurate method to assess more than one type of dexterity.

INTRODUCTION

Dexterity is usually defined as a function of control, the coordination of muscle movements usually in synchronisation with the eyes, and it can also be defined as the quality of motor skills of hands and fingers. It is a very valuable, versatile capacity involved in all variety of activities and situations, and demanded in many jobs and in everyday life. Dexterity in each person is qualitatively different and unique.

Dexterity has been assessed by occupational therapists using tests of hand skill as aptitude tests for selection and vocational guidance purposes, while physicians have used assessment methods to measure motor recovery of patients after stroke (Baum C, 1995) (Clark, Czaja, & Weber, 1990). Despite the frequent use of such tests, there is a notable lack of information regarding the definition of the factors involved in manipulative performance and the different types of dexterity that account for quality of function.

Although there has been plenty of work done on developing dexterity assessment methods, these methods have been designed for a specific application (rehabilitation, occupational therapy) and they are not suited for use for the evaluation of activities of daily living and accessibility of products (Yancosek & Howell, 2009) (Fleishman & Hempel, 1954) (Elfant, 1995) (Light C. , 2000) (Wolf, Lecraw, Barton, & Jann, 1989).

The development of assessment methods in other areas has been increasingly aimed at sampling as nearly as possible one ability category at a time; however, little research in this direction has been done when it comes to dexterity. In fact, dexterity is still defined by many researchers as if it were a unitary ability, even after many studies have found evidence indicating that this is not the case (Fleishman & Hempel, 1954) (Buxton, 1938).

The Variable Dexterity Test (VDT) was developed as part of a wider study, the broader aim being to fully understand dexterity and its affect on human-product interaction during activities of daily living. This is done with a view to the improvement of occupational therapy methods when assessing dexterity and general hand function. The proposed assessment method is based on the analysis of a wide range of grasping patterns through the time it takes a person to perform standard activities of daily living, each requiring different types of dextrous manipulation. For the purpose of this study, activities of daily living (ADL) are defined as daily self-care activities within an individual's place of residence, and they can be made of more than one particular task, e.g., grasping and manipulating shoelaces are only two of the activities required to put on shoes with laces.

Comment [VG1]: Reviewer's comment: "You have not said what you mean by this term in the paper (VDT), rather you used 4 actions that would form part of an activity of daily living (subtle difference). You express it better in Key Finding 1."

BACKGROUND

When trying to measure dexterity there are many tests and apparatus that have been developed and published to describe the quality of hand function and to quantify performance during manual activities (Tiffin, 1968) (Wolf, Lecraw, Barton, & Jann, 1989) (Moberg, 1958) (Light, Chappell, & Kyberd, 2002).

In order to accurately measure hand function one must take into account dynamics, perception of movement, speed of manipulation and their relationship to functional tasks, though, as mentioned above, many tests use time as the critical measure of performance.
A dexterity test is designed so that it provides information about the quality and the speed of performance of the hand while the subject accomplishes a task, making it easier for the examiner to quantify the quality of performance and manipulate the resulting data for deeper analysis.

Most tests make use of standardized objects such as pegs or blocks as test items to determine the person's dexterity while using their fingers to accomplish tasks. The Purdue Pegboard Test (Tiffin, 1968), Minnesota Manual Dexterity Test (Surrey, et al., 2003), Box and Block Test (Mathiowetz, Volland, Kashman, & Weber, 1985), O'Connor Dexterity Test (Hines & O'Connor, 1926), and, more recently, the Functional Dexterity Test (FDT) (Aaron & Stegink Jansen, 2003) have all been administrated as part of dexterity and hand function studies.

However, most available dexterity tests refer to dexterity as a unitary ability, even when previous works have demonstrated the existence of more than one type of dexterity and that each of these types is influenced by a unique dexterity factor (Fleishman & Hempel, 1954) (Napier, 1956).

The Functional Dexterity Test is an example of the few instruments aimed at assessing one specific type of grip and type of dexterity. It is focused on an individual prehension pattern, the 3-jaw chuck prehension, identified as the one humans use to perform common tasks such as writing, buttoning, and tying laces, combining thumb and fingers without the need of using the palm. The participant is asked to turn pegs starting from the top opposite corner from his dominant hand as quickly as possible without turning the hand up to face the ceiling or touching the board. Scores of other established tests and performance of ADL involving the 3-jaw-chuck prehension pattern (buttoning, tying shoe laces, screwing nut and bolt) were correlated with scores from the test in order to validate and prove the reliability of the method.

The tasks utilised by the FDT combine the three phases of dexterity: manipulation, time and accuracy. The test measures the ability of the subject to perform the grip pattern quantifying time during the manipulation of pegs in a pegboard comparing to performing daily precision tasks. Since the FDT was designed to assess one specific grasping pattern, it is useful to qualify the type of dexterity associated to that grip style.

Similarly, the Wolf Motor Test was developed to assess function of the upper extremity after a stroke, by analysing moving patterns ranging from simple movements (moving the hand from the lap to the table) to the manipulation of small objects (Wolf, Lecraw, Barton, & Jann, 1989). The tasks required by this test are mostly not-real-life tasks and the score is, the time it took the subject to complete the movements. Speed and quality of execution measurements demonstrated high reliability according to the work conducted by Richards et al (Reliability and validity of two tests of upper extremity motor function post-stroke, 2001).

The Arm Motor Mobility Test (AMAT) was designed to complement the Wolf Motor Test, adding more complex functional activities and evaluating quality of movement along with completion time (Kopp, et al., 1997). In another example of direct task evaluation, Desrosiers created the TEMPA (Test Evaluant les Membres Supériors de Personnes Âgées) assessing the performance of functional activities in a more standardised experimentation environment. The TEMPA measures time of execution as well as quality of movements. The functional ability rating is based on how much of the task was accomplished without the need of any adaptation and the scale intends to reflect the quality of motor components (0 = task completed without hesitation or any kind of adaptation, to -3 = task could not be performed beyond 25%) (Desrosiers, Hébert, Dutil, & Bravo, 1993).

Speeds of execution and functional ratings have demonstrated excellent reliability in previous studies, although some tasks have more reliability than others (Richards, Stoker-Yates, Pohl, Wallace, & Duncan, 2001).

Colin Light, Paul Chappell and Peter Kyberd developed the Southampton Hand Assessment Procedure (SHAP) in 2000 at the University of Southampton (Light C. , 2000). Aiming at the assessment of effectiveness of upper limb prostheses, the SHAP is now applied to general assessments of unimpaired participants. The SHAP consists of 8 abstract objects and 14 Activities of Daily Living. The participant times each task, avoiding reliability on the reaction times of the examiner. The test is performed with objects placed on a two-sided board.

Using the method described by Light, Chappell & Kyberd, the examiner can generate a quantifiable assessment of hand function that can be broken down into six prehensile patterns.

The examiner can visualise whether a participant has exceptional power and spherical grip, or impaired function in the ability to perform finer manipulations such as tip and tripod grips. A SHAP Index of Function score is generated, which is one number that provides an overall assessment of hand function.

It is worth noting that current dexterity tests have been designed for a specific application (e.g. rehabilitation, occupational therapy) and, although they have been shown to be reliable methods to assess hand function when used within their field of application, a method for the evaluation of function during activities of daily living looking towards a qualification of the accessibility of products is still needed (Yancosek & Howell, 2009) (Fleishman & Hempel, 1954) (Elfant, 1995) (Light C. , 2000).

In the current study, an effort in this direction is made, developing a method that samples each type of dexterity as purely as possible, making use of previous findings and recommendations toward this objective. The new dexterity test apparatus was developed aiming at a selection of different dexterous hand and finger motions in an effort to achieve the goal of individual dexterity type assessment through time and observed quality of performance.

Purpose

The purpose of this study is the introduction of a dexterity assessment method that is suitable for occupational therapy use, providing equipment construction standards and administration instructions. Reliability and validity of the newly developed method were studied in a pilot group and the methods and results are explained and discussed as well.

Method

Equipment construction Standards

The VDT is made of four subtests each named after the grasping pattern required for the completion of the task: Precision, Cylinder, Spherical and Extended Spherical.

The first three subtests are performed on a 270mm x 420mm and 15mm deep square wooden board with 8 holes. Each of the holes measures 15mm in depth and 74mm in diameter. Holes are separated from each other by 40mm in 2 rows of 4 holes each (Figure 1). For the VDT-Precision subtest there are 8 plastic objects made of a 73mm circular base and an attachable handle that is 30mm high and 70mm in length (Figure 2-A). The board and the pegs are sanded to a smooth surface.
The Cylinder subtest uses the same 8 circular bases with 8 attachable plastic cylinder-shaped handles. The plastic cylinders are 80mm tall and 40mm in diameter (Figure 2-B). The Spherical subtest's objects consist of the circular bases without any attachable handle.

FIGURE 1 SPHERICAL SUBTEST SETUP

A second wooden board is used for the 113mm subtest. The dimensions of the board are 440mm x 365mm, and 15mm deep. The board has 6 holes. Each of the holes measures 15mm in depth and 113mm in diameter. Holes are separated from each other by 44mm in 2 rows of 3.
The Extended Spherical subtest uses 6 circular-shaped plastic objects. The objects are 112mm in diameter (Figure 2-C).

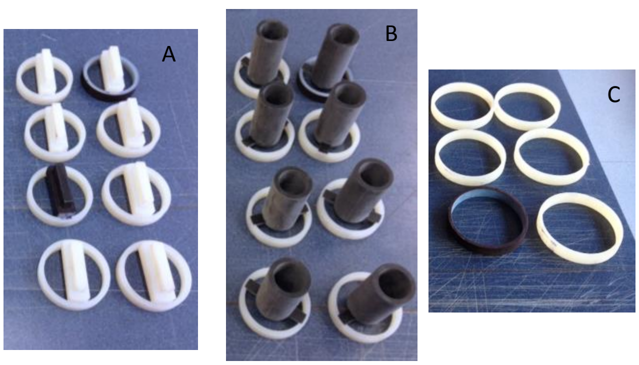

FIGURE 2 A) EXTENDED SPHERICAL SUBTEST OBJECTS; B) PRECISION SUBTEST OBJECTS; C) CYLINDRICAL SUBTEST OBJECTS

ADMINISTRATION OF THE VDT

The board corresponding to the current subtest is placed 10 cm from the edge of a table where the participant is comfortably sitting. The test objects are placed aligned 10 cm to the side of the board corresponding to the participant's dominant hand.
The participant is instructed to pick up one object at a time with the dominant hand, starting with the object at the top (away from participant) and place it in a hole of the board starting at the top-opposite side of the board. The participant will continue until she or he has put all the objects onto the board (Figure 3).

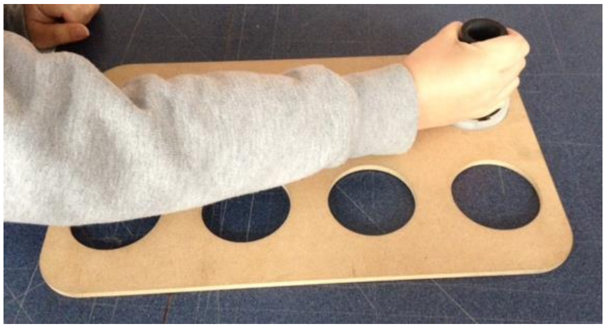

FIGURE 3 PARTICIPANT PERFORMING THE VDT-CYLINDER SUBTEST

The following verbal instructions should be provided to the participant: "Please start with your dominant hand. Start by picking up the object at the top [point to object] and place it in the hole at top opposite side of the board, pick and place all the objects as quickly as possible, finishing with the object at the bottom of the line. If you drop an object, time is stopped, and a 5-second penalty is added. Continue to pick and place the objects with the object that you just dropped. The clock starts where it was stopped, and the time is continued."

The examiner demonstrates each subtest by doing 2 objects. The participant is asked to practice by doing the test one time. The test is then performed once for each type of object (subtests). For each type of object, the examiner records the time it takes complete the test, penalties and unusual movement patterns observed.

SCORING

The score is the time it takes the participant to put all the objects on the board. For every time a participant drops an object, time is stopped, and a 5-second penalty is added.
Both penalized and non-penalised times are registered.
The examiner should note any unusual movement during completion of the test as well as a description of the main grasping pattern features for each subtest. The final score, time plus penalties is the level of dexterity for each specific grip style.
Time, penalties, final score and notes regarding movement patterns should all be included in the assessment of the participant's dexterity performance for each of the types of dexterity under analysis. Further normative studies will provide examiners with an interval level of measurement for parametric statistical analysis of data.

RELIABILITY STUDY

The extent to which a measurement is consistent and free from error is called reliability. Reliability estimates were obtained involving a pilot group of 24 healthy participants (12 female, mixed backgrounds: 8 Asian, 12 European, 4 Latin American). The purpose of the study was to demonstrate that each subtest is truly objective. The test must produce consistent results among the pilot group with both a single assessor (test-retest reliability) and multiple assessors (inter-rater reliability).

Comment [VG2]: Reviewer's comment: "Sample: there is insufficient detail of those in the sample. It would be useful to know country of origin of the sample (British/European, Americas, Asia etc) and more details"

A first rater assessed one control group (24 healthy participants, 13 males ranging in age from 20 to 50 years and 11 females ranging in age from 22 to 45 years), with 3 replicate evaluations for each subject. To establish test-retest reliability, it is necessary to show there is no statistically

significant effect in the replicate trials. The most appropriate method to determine whether the data have test-retest reliability is by an analysis of variance (ANOVA) with the null hypothesis being that no significant difference exists between replicates. The null hypothesis was tested at a p level of 0.05.
Inter-rater reliability was assessed for the same pilot group with 3 raters administrating the VDT.

The F value obtained from the ANOVA must exceed a critical value (3.13), which is based on the 95% confidence interval in order to prove that the null hypothesis should be rejected and, therefore, that there is a statistically significant difference between replicates. The analysis of variance was performed for all subjects and replicates on a subtest-by-subtest basis under the assumption that if all subtests are repeatable, then the complete assessment can be considered as reliable.

VALIDITY STUDY

A relationship between scores on the VDT and performance of activities of daily living has to be determined in order to know if the method is measuring what it is intended to measure.
The VDT was administered to the control group, and the scores were correlated with the participant's ability to perform 4 activities of daily living—opening a soft drink bottle, tying shoelaces, opening a jar, and buttoning a shirt- as representative of activities requiring a variety of grasping patterns.
An examiner recorded the time of completion of each of the ADL and Bi-serial Correlation was performed to determine whether a relationship exists between each of the 4 subtests of the VDT and the activities.
Validity can also be established by correlating scores on the new test with scores of an established test, known as the gold standard. The relationships between scores from the VDT and the Purdue Pegboard Test were evaluated for the same control group. The Purdue Pegboard Test was chosen as the standard because of its long history of usage and reliability among researchers.

FINDINGS

RELIABILITY STUDY

Analysis of variance indicated that no significant differences existed between replicates on participants who received instructions from one instructor (test-retest). The analysis of variance was performed for all subjects and replicates on a subtest-by-subtest basis under the assumption that if a procedure is repeatable, then the complete assessment can be considered as reliable. The result indicates that the F critical value was not exceeded for any of the 4 subtests (Table 1).

	Sum of Squares	df	Mean Square	F	Sig.
VDT-Spherical					
Between Groups	7.441	2	3.721	2.136	.126
Within Groups	120.172	69	1.742		
Total	127.613	71			
VDT-Extended spherical					
Between Groups	7.370	2	3.685	1.027	.363
Within Groups	247.520	69	3.587		
Total	254.890	71			
VDT-Precision					
Between Groups	10.242	2	5.121	1.750	.181
Within Groups	201.878	69	2.926		
Total	212.120	71			
VDT-Cylinder					
Between Groups	10.901	2	5.451	2.089	.132
Within Groups	180.017	69	2.609		
Total	190.918	71			

Table 1 Test-Retest ANOVA score results for each VDT subtest

The ANOVA test revealed an F maximum value of 2.136 (F critical value 3.13) and *P* minimum value of 0.126 for the VDT-Spherical subtest, indicating that, although there is no statistically adverse effect in the repeatability of any of the subtests, this subtest is less repeatable than the others.

The inter-rater ANOVA test showed statistically insignificant effect for all subtests but one. The VDT-Spherical was demonstrated to have significant differences between raters (F = 3.601, F critical value = 3.13, P = 0.033) at the 95% confidence interval level. For the other three subtests the examiner appears to have statistically little effect on the performance of the test, with a maximum F value of 1.653 and minimum P value of 0.199 for the VDT-Cylinder subtest, thereby indicating inter-rater reliability (Table 2).

	Sum of Squares	df	Mean Square	F	Sig.
VDT-Spherical					
Between Groups	7.614	2	3.807	3.601	.033
Within Groups	72.938	69	1.057		
Total	80.552	71			
VDT-Extended spherical					
Between Groups	8.146	2	4.073	1.347	.267
Within Groups	208.597	69	3.023		
Total	216.743	71			
VDT-Precision					
Between Groups	6.683	2	3.342	1.007	.370
Within Groups	228.874	69	3.317		
Total	235.557	71			
VDT-Cylinder					
Between Groups	8.938	2	4.469	1.653	.199
Within Groups	186.525	69	2.703		
Total	195.463	71			

TABLE 2 INTER-RATER ANOVA SCORE RESULTS FOR EACH OF THE VDT SUBTESTS

VALIDITY STUDY

Correlation between precision activities and the scores on the VDT-Precision subtest was 0.644 for buttoning and 0.603 for shoelaces, indicating there is correlation between performance of this subtest and activities requiring finer dexterity (Significance at the 0.01 level, 2-tailed).

The VDT-Cylinder and VDT-Spherical, subtests showed correlation with activities requiring grosser dexterity (power, cylinder and spherical grip styles), with the highest correlation being those found between VDT-Cylinder and Opening a soft drink bottle (0.707) and between VDT-Spherical and Opening a soft drink bottle (0.646) (Table 3).

	VDT Precision/ Tying Shoelaces	**VDT Precision/ Buttoning a shirt**	**VDT Cylinder/ Opening soft drink bottle**	**VDT Spherical/ Opening soft drink bottle**
Pearson Correlation	0.603	0.644	0.707	0.646
Significance (2-tailed)	0.002	0.001	0.0001	0.001
N	24	24	24	24

TABLE 3 CORRELATIONS BETWEEN VDT SUBTESTS AND ACTIITIES OF DAILY LIVING

The Pearson Bi-serial correlation between the scores on the Purdue Pegboard Test Right Hand and those subtests of the VDT that require finer dexterity was high (-0.813 for the VDT-Spherical, -0.849 for the VDT-Precision, and -0.617 for the VDT Cylinder subtest) indicating the VDT selected subtests results agree with those coming from a well established dexterity test (Table 4).

	VDT Precision/PPBT	VDT Spherical/ PPBT	VDT Cylinder/PPBT
Pearson Correlation	0.849	0.813	0.617
Significance (2-tailed)	0.0001	0.001	0.001
N	24	24	24

TABLE 4 CORRELATIONS BETWEEN VDT SUBTESTS AND PURDUE PEGBOARD TEST

DISCUSSION

Comment [VG3]: Changed almost completely. Reviewer's comment: "Discussion: is very short and does not really discuss the findings, and it only focuses on future studies"

For the pilot group used for this study the VDT proved to be a flexible and reliable tool that approaches the problem of assessing dexterity with a more robust method.

It was shown, through analysis of variance, that the VDT is repeatable between replicates on participants who received instructions from one instructor (the F critical value was not exceeded for any of the 4 subtests), although it was evident the existence of a learning effect that diminished greatly after the second trial.

Through ANOVA, the Spherical subtest proved to be the less repeatable of the VDT subtests, indicating that the motions and grasping patterns required to complete that specific subtest are more prone to be affected by the learning effect and more work will be done to improve in this regard.

Inter-rater reliability from the VDT-Spherical proved to be significantly lower ($F = 3.601$, F critical value = 3.13, $P = 0.033$) than that from the other three subtests, further indicating that this subtest is one of the most important points to be improved from the method.

The VDT subtests scores (times) were correlated with times to complete activities of daily living selected according to the grasping patterns and the type of dexterity that each subtest was designed to assess, with the VDT-Precision subtest showing strong correlation with a selection of activities of daily living that require finer, more precise hand coordination. The highest correlation was found to be that between the VDT-Cylinder and the opening of a soft drink bottle (0.707), indicating that the quality of performance of that activity is highly influenced by the cylinder grasping pattern.

These results also indicate that the quickest the participant was able to perform the ADL, the less time it took to complete each of the subtests that make the VDT.

Similarly, it was shown, through Pearson Bi-serial correlation between the scores on the Purdue Pegboard Test Right Hand and the subtests of the VDT that require finer hand manipulation, that the proposed method delivers results that are consistent with those delivered by a well established dexterity test (-0.849 for the VDT-Precision). Subtests requiring grosser types of dexterity (0.617 for the VDT-Cylinder) showed lower levels of correlation with the Purdue Pegboard Test, indicating that they are not assessing the exact same function. These grosser subtests will be compared with other gold-standard tests in the future aiming at the complete validation of the VDT.

It was shown that the method matches a wider range of grasping patterns, and it was demonstrated that it effectively measures different types of dexterity in relation to actual activities of daily living.

The VDT has the potential to directly impact hand rehabilitation practice, providing therapists with a robust, easy to administer tool for the assessment of hand function.

Studies involving a larger number of participants over a wider age range is needed to generate data that include mean values and confidence intervals to define which boundaries of scores can be considered to be low levels of dexterity.

Comment [VG4]: Reviewer's comment: "Make sure your claims are supported by evidence in the paper"

CONCLUSION

The VDT construction standards and design make the apparatus an easy to administer solution to the problem of assessing dexterity when compared to existing dexterity tests.

The Variable Dexterity Test produces consistent results among a pilot group with both a single assessor (test-retest reliability) and multiple assessors (inter-rater reliability).

Strong correlation between scores from the VDT and performance of activities of daily living has been proved, showing that the method is measuring what it is intended to measure.

The results indicate that the better the participant was able to perform the ADL, the less time it took to complete each of the subtests that make the VDT.

Validity was also established by correlating scores on the new test with scores of an established, well known test, the Purdue Pegboard Test.

The Pearson Bi-serial correlation between the scores on the Purdue Pegboard Test and those subtests of the VDT that require finer dexterity was shown to be strong.

KEY FINDINGS

- The VDT measures a range of types of dexterity that are required to perform activities of daily living (ADL).
- The VDT was shown to be reliable via ANOVA in Test-retest and Inter-rater trials for a pilot control group.
- The VDT was shown to be valid by correlating scores from the test with proficiency to complete ADL and a gold-standard dexterity test.

WHAT THE STUDY HAS ADDED

This study presents a new dexterity assessment method, capable of measure individual types of dexterity that are required for the completion of activities of daily living.

Acknowledgments

This work was supported by CONACYT (Consejo Nacional de Ciencia y Tecnología), Mexico's official entity in charge of the promotion of scientific and technological activities. The helpful suggestions of Claudia Mazza are gratefully acknowledged.

References

Aaron, D. H., & Stegink Jansen, C. W. (2003). Development of the Functional Dexterity Test (FDT): Construction, Validity, Reliability, and Normative Data. *Journal of Hand Therapy* .

Baum C, E. D. (1995). Occupational performance: occupational therapy's definitions of function. *Am J Occup Ther. , 49*, 1019-1020.

Bernstein, N. A. (1996). *Dexterity and its development.* Mahwah, NJ: Lawrence Erlbaum Associates.

Buxton, C. E. (1938). The application of factorial methods to the study of motor abilities. *Psychometrika , 3*, 85-93.

Clark, C. M., Czaja, S. J., & Weber, R. A. (1990). Older adults and daily living task profiles. HumanFactors. *Human Factors , 5* (32), 537-549.

Desrosiers, J., Hébert, R., Dutil, E., & Bravo, G. (1993). Development and reliability of an upper extremity funtion test for the elderly: The TEMPA. *Canadian Journal of Occupational Therapy , 60*, 9-16.

Elfant, I. L. (1995). Correlation between kinesthetic discrimination and manual dexterity. *Am J of Occup Th , 31*, 23-28.

Fleishman, E. A., & Hempel, W. E. (1954). A Factor Analysis of Dexterity Tests. *Personnel Psychology , 7* (1), 15-32.

Hines, M., & O'Connor, J. (1926). A measure of finger dexterity. *J. Pers. Res.* , 379-382.

Kopp, B., Kunkel, A., Flor, H., Platz, T., Rose, U., Mauritz, K., et al. (1997). The Arm Motor Ability Test: Reliability, validity and sensitivity to change of an instrument for assessing disabilities in activities of daily living. *Archives of physical medicine and rehabilitation , 78*, 615-620.

Light, C. M., Chappell, P. H., & Kyberd, P. J. (2002). Establishing a Standarized Clinical Assessment Tool of Pathologic and Prosthetic Hand Function: Normative Data, Reliability, and Validity. *Arch Phys Med Rehabil* .

Light, C. (2000). Southampton Hand Assessment Procedure: User Manual. *Department of Electronics and Computer Science, University of Southampton.*

Mathiowetz, V., Volland, G., Kashman, N., & Weber, K. (1985). Adult Norms for the Box and Block Test of Manual Dexterity. *The American Journal of Occupational Therapy* .

McPhee, S. D. (1987). Functional hand evaluations: a review. *Am J Occup Ther. , 41*, 158-163.

McWilliam, R. P. (1970). A list of everyday tasks for use in prosthesis design and development. *Bull Prosthet Res* , 135-64.

Moberg, E. (1958). Objective methods for determining the functional value of sensibility in the hand. *J Bone Joint Surg. , 40B*, 454-476.

Morrey , B. F., Askew, L. J., An, K. N., & Chao, E. Y. (1981). A biomechanical study of normal functional elbow motion. *J. Bone Jt Surg., 1981,63A(6),872-877 , 6* (63), 872-877.

Murgia, A., Kyberd, P. J., Chappell, P. H., & Light, C. M. (2004). Marker placement to describe the wrist movements during activities of daily living in cyclical tasks. *Clinical Biomechanics , 19*, 248–254.

Musur-Grieve, M. (1984). Methods of assessment and management of the rheumatoid hand at the Institute of Rheumatology, Wasaw, Poland. In J. M. Hunter, & e. al, *Rehabilitation of the Hand* (p. 651). St. Louis, Mo.: Mosby.

Napier, J. R. (1956). The prehensile movements of the human hand. *J. Bone Joint Surg. , 38B*, 902-913.

Parry, C. W. (1966). *Rehabilitation of the Hand* (Second Edition ed.). London: Buttenworths.

Richards, L., Stoker-Yates, J., Pohl, P., Wallace, D., & Duncan, P. (2001). Reliability and validity of two tests of upper extremity motor function post-stroke. *The Occupational Therapy Journal of Research , 21* (3), 201.

Roeder, W. S. (1958). Manipulative Aptitude Test Apparatus. *United States Patent Office 2,835,986* .

Rowson, J., Sangrar, A., Rodriguez-Falcon, E., Bell, A. F., Walton, K. A., & Yoxall, A. (2011). Rating accessibility of packaging: A medical packaging example.

Stammers, R. B. (1995). Factors limiting the development of task analysis. *Ergonomics , 3* (38), 588-594.

Stanton, N. A., & Barber, C. (2005). Validating task analysis for error identification: reliability and validity of a human error prediction technique. *Ergonomics , 48* (9), 1097-1113.

Surrey, L. R., Nelson, K., Delelio, C., Mathie-Majors, D., Omel-Edwards, N., Shumaker, J., et al. (2003). A comparison of performance outcomes between the Minnesota Rate of Manipulation Test and the Minnesota Manual Dexterity Test. *Work* , 97-102.

Taylor, C. L., & Schwarz, R. J. (1955). The Anatomy and Mechanics of the Human Hand. *Artificial limbs , 2* (2), 22-35.

Tenneti, R., Johnson, D., Goldenberg, L., Parker, R. A., & Huppert, F. A. (2012). Towards a capabilities database to inform inclusive design: Experimental investigation of effective survey-based predictors of human-product interaction. *Applied Ergonomics* , 713-726.

Thurstone, L. L. (1955). *The Vectors of Mind.* Chicago, Illinois: University of Chicago.

Tiffin, J. (1968). Purdue Pegboard Test.

Trombly, C. A., & Scott, A. D. (1989). Evaluation of motor control. In C. A. Trombly, *Occupational Therapy for Physical Dysfunction* (pp. 55-71). Baltimore, Md: Williams & Wilkins.

Valero-Cuevas , F. J., Smaby, N., Venkadesan, M., Peterson, M., & Wright, T. (2003). The strength–dexterity test as a measure of dynamic pinch performance. *Journal of Biomechanics , 36*, 265-270.

Wolf, S., Lecraw, D., Barton, L., & Jann, B. (1989). Forced use of hemiplegic upper extremities to reverse the effect of learned nonuse among chronic stroke and head-injured patients. *Experimental Neurology , 104*, 125-132.

Yancosek, K. E., & Howell, D. (2009). A Narrative Review of Dexterity Assessments. *Journal of Hand Therapy , 22* (3), 258-69.