The Open Journal of Occupational Therapy

Volume 3
Issue 3 *Summer 2015*

Article 5

7-1-2015

Examining the Purdue Pegboard Test for Occupational Therapy Practice

Debra K. Lindstrom-Hazel
Western Michigan University, debra.lindstrom@wmich.edu

Nicole VanderVlies Veenstra
Mary Free Bed Rehabiliation Center, n3vander@gmail.com

Recommended Citation
Lindstrom-Hazel, D. K., & VanderVlies Veenstra, N. (2015). Examining the Purdue Pegboard Test for Occupational Therapy Practice. *The Open Journal of Occupational Therapy, 3*(3). https://doi.org/10.15453/2168-6408.1178

Examining the Purdue Pegboard Test for Occupational Therapy Practice

Abstract

Background: Occupational therapy ethics require that therapists use current assessment tools that provide useful comparison data. When an assessment only has normative data that is more than 40 years old, it cannot be considered current. The purpose of this study was to examine the past and current use of the Purdue Pegboard Test by occupational therapists and other professionals and to determine if it is beneficial to conduct a large normative study on the Purdue Pegboard Assembly Task (PPAT) in order to bring the test up to date.

Method: This was a psychometric study of inter-rater reliability and a small normative study of the PPAT with 150 healthy working adults from MI. Descriptive statistics were used for normative means, standard deviations, and standard errors of measurement.

Results: Inter-rater reliability was measured using the intra-class correlation coefficient for the mean of all student-rating teams of seven occupational therapy students. The result of the psychometric study determined the ICC was above .99. During the normative study, 150 participants performed the PPAT for three trials. Norms for gender and ages 18-49 and 50-62 are presented.

Conclusion: The result of the inter-rater reliability test determined that OT students can be reliable raters for the PPAT. The normative study collected current norms for healthy working adults in MI, but validity testing and a larger normative study is needed to bring the psychometrics of the PPAT up to date to be generalized for current use by occupational therapists.

Keywords

fine motor, dexterity, psychometrics, inter-rater reliability

Cover Page Footnote

The authors would like to thank Samantha Harrow Swantek and Erin Courtnay Catalano for their participation in the planning, Inter-Rater Reliability Scoring and Administration Instructions, and data collection. They also wish to thank Ashley Thompson for assistance with data analysis and Emily Byl and Tifanie Rider for editorial assistance.

Credentials Display

Debra Lindstrom-Hazel, Ph.D., OTRL, FAOTA

Nicole VanderVlies Veenstra, MS, OTRL

DOI: 10.15453/2168-6408.1178

As occupational therapy practitioners strive to provide best practice and evidence-based therapy, the first decision they make for any client is how to evaluate the person to determine his or her occupational needs. The occupational therapist (OT) needs to determine which assessment instrument will provide the most meaningful measurement of the client's status that is based on good, current evidence (Ottenbacher, Tickle-Degnen, & Hasselkus, 2002). The OT's evaluation results need to guide intervention and recommendations as well as measure the client's progress and intervention outcomes. Following *The Occupational Therapy Code of Ethics and Ethics Standards* (American Occupational Therapy Association [AOTA], 2010), OTs must avoid "the inappropriate use of outdated or obsolete tests/assessments or data obtained from such tests" (p. S19). Unfortunately, some of the assessments that the profession considered the "gold standard" 30 years ago have become outdated in the years since they were developed and/or last revised. An OT may need to choose an assessment that may not measure exactly the concept he or she needs to measure because the assessment that the OT considers to have more face validity has outdated norms. This can create a void in the battery of instruments that are available to an OT. There is a strong call in the occupational therapy literature to center assessments and interventions directly on the specific occupations rather than on simulated skills (Fisher, 2013), but the present clinical reality in the United States is that there are times when an OT needs to administer a component-based assessment at a clinical setting and make occupationally based recommendations accordingly. An OT may need to make recommendations for a client's ability to return to work after a hand injury when the essential job functions of the work require finger dexterity and the specific job task required on the job cannot be performed in the clinic.

Literature Review

In a systematic review by Causby, Reed, McDonnell, and Hillier (2014), the Purdue Pegboard Test (PPT) was identified as one of the top three assessments of hand dexterity for health care professionals, due to its relatively higher reliability and validity and fewer confounding variables, such as age, gender, and handedness. Although the authors recommended the PPT as an instrument to use with health care professionals, they did call for a more rigorous evaluation of its validity.

Despite its high rating in this systematic review it is not one of the assessments included in the National Institutes of Health's (NIH) Toolbox, which was recently published with updated norms. The Toolbox meets the NIH goal of a standard set of brief measures that many professionals can use to assess and compare cognitive, emotional, motor, and sensory function for people 3 to 85 years of age. The Nine Hole Peg Test (NHPT) was the only dexterity assessment that was included in the Toolbox (NIH, 2012).

Although the NHPT has good clinical validity for use by OTs in some settings, the PPT has been found to require more fine motor precision and to be more sensitive for detecting functional impairment in young and middle-aged subjects (Amirjani, Ashworth, Olson, Morhart, & Chan,

2011; Mathiowetz, Rogers, Dowe-Keval, Donahoe, & Rennells, 1986). Fleishman and Ellison (1962) factored out the fine motor dexterity used in the PPT from other types of hand dexterity involved in determining workers' ability to do tasks, such as assemble small parts and wire electrical circuits. This fine motor or finger dexterity could also be needed for avocational skills, such as baiting a hook or making jewelry. Strauss, Sherman, & Spreen (2006) found that the PPT demanded more cognitive speed and attention control than other dexterity tests, making it relevant for predicting complex dexterity function in daily life.

The PPT is a standardized assessment that uses four different subtests to determine the user's level of dexterity. Tiffin and Asher (1948) developed the evaluation in 1948 to assess functional dexterity among factory workers. The researchers created specific norms based on those who worked common jobs of that time period, such as assembling small parts (factory workers), manipulating small and complex items (sewing machine operator), typing on a manual typewriter and writing by hand (college students), and prior military experience (veterans). Each category of workers was normed separately. Tiffin and Asher had a large sample of more than 7,814 subjects, both male and female, aged 18 years and over. In 1968, Tiffin revised the Purdue Pegboard Examiner's Manual to update the norms in the same categories as the 1948 norms. The categories included reflect different occupations than the occupations that people are involved in today. Two categories in those norms are college students and military personnel. Both of these categories contain examples of how different the occupations are in 2015 from 1948 and 1968. The way people type today is different than the way people typed in 1948 and 1968 (manual typewriters vs. computer/laptop keyboards), and the use of touch technology (smartphones and small hand-held devices) has changed the way college students use their hands. Military personnel are trained in a different manner and require a variety of diverse skills that have evolved since the 1940s and 1960s. In addition, those who work in industrial settings may use computers frequently, even when performing small piece assembly. Since the revision of the manual in 1968, the validity and reliability of this assessment has been tested on numerous occasions with several different populations (Amirjani et al., 2011; Gallus & Mathiowetz, 2003; Mathiowetz et al., 1986), but new norms for the general population have not been published.

More recent norms have been developed for specific populations, including people between 14 and 19 years of age (Mathiowetz et al., 1986), aged 40 years and over (Agnew, Bolla-Wilson, Kawas, & Bleecker, 1988), aged 60 years and over (Desrosiers, Hébert, Bravo, & Dutil, 1995), and people with multiple sclerosis (Gallus & Mathiowetz, 2003). The 1968 norms are currently the only available norms that include everyone who could be compared to a healthy employed population today. Current general adult norms are also needed for people who experience temporary injuries, illnesses, or diminished function to understand the extent of their limitations for finger dexterity and to document progress after therapy

DOI: 10.15453/2168-6408.1178

intervention when this needs to be conducted in a clinical setting.

Use of the PPT in Occupational Therapy Practice

An industrial psychologist created the PPT to assess assembly line workers' dexterity (Tifflin & Asher, 1948), and neuropsychologists and psychologists have used the PPT as part of a battery to assess manual dexterity and bimanual coordination as part of neuropsychological testing (Strauss et al., 2006). The first evidence in the *American Journal of Occupational Therapy* (AJOT) literature related to OTs using the PPT in practice was in 1986 (Mathiowetz et al., 1986). At that time, the PPT was reported to be in use in vocational training as a fine motor assessment. According to an OT at a state vocational training school, the PPT is no longer used because the 1968 norms were not valid when compared with the OT's observation of the client's functional performance (R. Lyon, personal communication, February 13, 2014).

The PPT is currently listed as a potential assessment in Doucet, Woodson, & Watford's centennial vision for rehabilitation intervention research (2014) as one option for assessing fine motor/finger dexterity. The four subtests in the PPT involve timing a subject's ability to place pegs in small holes with his or her dominant hand, his or her non-dominant hand, with both hands simultaneously, and then a bilateral assembly task (PPAT). The standardized administration for the PPAT has the examiner explain, demonstrate, and allow the person to practice the specific combination of a peg, a collar, and two washers in specific pegboard holes (Tiffin & Asher, 1948). Although it is not exactly like a task that employees may describe for their vocation/avocation, the PPAT is a complex finger dexterity task that can be evaluated in a clinical setting in a short period of time.

PPT Psychometrics

In a review of the psychometric values of 14 different dexterity evaluations, Yancosek and Howell (2009) found that the PPT had high validity and reliability based on its initial use with a healthy population and its later use with populations with limitations (i.e., carpal tunnel syndrome, multiple sclerosis). Yancosek and Howell asserted that three-trial administration was more reliable than the one-trial administration. However, according to Gallus and Mathiowetz (2003), a one-trial administration was sufficient for use with clients who have multiple sclerosis.

In 2011, Amirjani et al. examined the reliability and validity of the PPT for people with carpal tunnel syndrome (CTS). This study included 190 subjects with CTS and 122 healthy subjects. The results of this study indicated that the PPT is considered a useful outcomes measure for people with CTS, as well as for healthy young and middle-aged adults. The 0.91 intra-class correlation coefficient (ICC) of the raters substantiated this interpretation (Amirjani et al., 2011). An ICC is typically used when comparing the results of two or more raters since the ICC includes a calculation of the reliability index of the measurement error between judges. ICCs have been reported to be the best method for reliability analysis (Buddenberg & Davis, 2000). Lee et al. (2013) found that the PPT had moderate-to-good test-retest reliability rating

for individuals who have schizophrenia; however, the measure of random error was considered substantial with this population. Thus, clinicians should be aware that random error may occur when using the PPT with people with schizophrenia.

Buddenberg and Davis (2000) provided evidence that the test-retest reliability of the PPT for the three-trial administration was better than the one-trial administration. These researchers provided evidence which demonstrated excellent reliability with the PPT as each of the three-trial administration correlations had an ICC > .80; the one-trial administration yielded an ICC of < .71. However, in contrast, Gallus and Mathiowetz (2003) examined the test-retest reliability of the PPT and did not discover any significant differences between one trial and the average of three trials for use with people who have multiple sclerosis. Thus, Gallus and Mathiowetz (2003) concluded that one-trial administration was sufficient for the people with multiple sclerosis in this study. As there are currently numerous studies that support the reliability and validity of the PPT (Amirjani et al., 2011; Buddenberg & Davis, 2000; Gallus & Mathiowetz, 2003; Lee et al., 2013), the next step in the process of updating the test is collecting updated, accurate norms for the populations with which this evaluation is used (Buddenberg & Davis, 2000).

One potential reason for the old norms could be related to the difficulty in conducting norm studies. One of the problems in getting large enough numbers for a reasonable study is the need to have multiple raters. When there are multiple raters, there needs to be excellent inter-rater reliability (IRR) demonstrated to maximize the validity of the assessment and norms. Occupational therapy students have been found to be reliable raters after training in collecting grip and pinch measurements (Lindstrom-Hazel, Kratt, & Bix, 2009). With careful attention to training and the use of standardized administration guidelines, research assistants with no previous background in health care administered various occupational therapy assessments, including the Barthel Index, the Mini Mental State Examination, the Philadelphia Geriatric Center Morale Scale, and a scale of instrumental activities of daily living with excellent reliability (Edwards, Feightner, & Goldsmith, 1995). From these past results, it is reasonable to believe that norms could potentially be collected for the PPT by multiple raters if good training materials are provided for administration and scoring.

Researchers designed the PPT to predict who would succeed in specific types of employment settings (Tiffin, 1968). However, the types of work that people do and current job requirements have changed in many ways since the 1980s (Phillips, Lindstrom-Hazel, Harrow Swantek, & Courtnay Catalano, 2013). Researchers have recommended that norms for instruments like the PPT be updated every 15-20 years (Strauss et al., 2006). Since the norms for the PPT are significantly outdated, the purpose of this study was to develop updated normative data for the Assembly Task of the PPT in healthy working adults. The researchers in this study only selected to study the PPAT because it is an assessment that focuses on bilateral hand use and is most representative of the types of finger dexterity movements that are involved in the

DOI: 10.15453/2168-6408.1178

occupations that people perform in their everyday lives. There is a need for an easy and quick to administer finger dexterity test when it is not possible to actually evaluate the person performing the occupational tasks they want or need to complete.

Methodology

This study included two parts for the PPAT: (1) a psychometric study that included expanded administration and scoring instructions and an IRR study with seven student raters, and (2) a normative study of the PPAT for healthy working adults. This data collection was one part of a larger study that also included data collection for grip and pinch strength norms (Phillips et al., 2013). Prior to beginning the studies, the researchers received HSIRB approval for both parts of this study.

Inter-Rater Reliability Study

The IRR study took place at a midwestern university dining hall where over 100 students, staff, and faculty voluntarily participated. The purpose of this study was to compare scores between the student investigators' recorded scores; it was not to examine the participants' individual performance.

Student raters and training. A convenience sample of seven student investigators and one supervising student investigator, all of whom were occupational therapy students at an accredited midwestern university, participated in this study. Five of the eight students had finished their occupational therapy preparation courses and were ready to begin a Level II fieldwork experience. Three of the students, including the supervising student investigator, were graduate-level students who had completed at least one Level II fieldwork experience. To ensure that the data collected by student investigators was accurate and consistent, expanded administration directions were written and the student data collectors participated in an IRR training session in which they scored pictures of the tasks that were displayed on a computer. Each student data collector completed the Purdue Inter-Rater Reliability Training Form for the 10 training pictures. The supervising student investigator scored the students' forms and checked them for accuracy with the principal investigator. This served as a competency test for scoring. The researchers reviewed any scoring errors with the student investigator, and offered individualized assistance until the principal investigator was confident that the student investigator was an accurate scorer. Only raters who had passed the competency test were allowed to collect data in this study.

The seven student data collectors were placed in groups of three for comparison rating. One student data collector administered three trials of the PPAT while the other two student data collectors were seated on either side of the administrator. Each of the three student data collectors scored the three trials completed by the research participants. Student data collectors collected data for periods of one hour following a "warm-up" session in which they reviewed the scoring procedures and rules from the initial training session. After one hour of data collection, all student data collectors took a fifteen-minute break before beginning another one-hour session of data collection (if they were scheduled for two data

collection sessions in one evening). Every student data collector was observed administering the PPAT for adherence to the administration protocol. The data was analyzed using Excel's Data Analysis Tool Pack using an intra-class correlation (ICC, Shrout & Fleiss, 1979).

Normative Study Participants

The sample for the normative study was a convenience sample of 150 research participants who were recruited from employees at car factories in southeast Michigan and health care organizations in western Michigan. The age range for the sample was between 18 and 62 years of age, and the average age was 49.15. Fifty-six percent of the research participants were male (n = 80) and 44% of the participants were female (n = 70). The research participants included office workers at both types of sites; factory workers at the car manufacturing plants; and health care, food service, and housekeeping workers at the health care sites. The participants were categorized into two age groups of 20 to 49 and 50 to 62 years of age and were classified by gender. Prior to engaging in the study, all of the research participants signed an informed consent form and stated that they were free from any conditions that limited their use of their upper extremities and did not have any work restrictions. An incentive of an apple or candy bar was available for all participants after completing the study.

Results

Inter-Rater Reliability of One versus Three Trials

The ICC were examined for the seven raters of the PPAT for the average of three trials for each participant. A confidence interval of 95% was used to obtain comparisons between the raters. Each student data collector scored a minimum of 64 participants' three trial completions.

The ICC score for all of the student-rating teams was above .97, with many of the scores above .99 ($p < .05$). Table 1 shows the separate, comparative ICC analysis between each student-rater. The results indicate that the average ICC of all rating pairs for the seven occupational therapy students was .99455, with each team averaging 13.9 scored subjects. See Table 2 for ICC averages of the ICC calculations for each of the teams of student raters.

Table 1

Individual Student Rater Comparisons

Raters	Subjects	ICC
1&2	10	1
1&3	11	0.999
4&5	12	1
1&6	11	0.996
1&7	11	1
2&4	35	0.995
2&6	14	0.981
3&5	22	0.981
4&7	9	0.998
3&6	11	0.987
5&7	11	0.987
2&3	11	0.999
1&4	11	0.999
5&1	10	0.999
6&7	11	0.998
4&6	14	0.982
2&7	20	0.999
4&3	11	1
5&6	11	0.992
3&7	22	0.999
Avg.	13.9	0.99455

DOI: 10.15453/2168-6408.1178

Table 2

Intra-Class Correlation Coefficients

	Intraclass Correlation	95% Confidence Interval	
		Lower Bound	Upper Bound
Single Measures	.995	.994	.996
Average Measures	.977	.997	.998

Participants

The researchers timed 150 Michigan residents between 18 and 62 years of age as they assembled as many sets of small pieces in the pegboard in the correct order as they could in one minute. They practiced the task once (according to the standardized protocol) and then were timed to complete the task three times. The scores were initially analyzed in 5-year age categories, for example, participants 41 to 45 years of age. The mean score differences in each age category were analyzed using a visual analysis of the graph, and it was determined that there were not enough participants in each of the categories and the means were not consistently different between the age categories for either males or females until 50 years of age and over for both males and females. The age categories were then divided into two categories, 18 to 49 and 50 to 62 years of age, to allow a reasonable number of participants in each category and to reflect when the mean scores started decreasing and the standard of error started increasing. Figure 1 shows the comparison of male and female means for the PPAT, and Figures 2 and 3 show the mean scores of the PPAT for 5-year age categories.

The standard error for females in the upper age category was 1.3 and the standard error for males in the upper age category was .80. In the younger age group of 18 to 49 years of age, the standard errors were low, with all falling below .62 for both gender groups with the standard errors for the males lower than the females. Table 3 provides the means, sample size (N), plus or minus 1 standard deviation (+- 1SD) and standard error (SE) and age and gender for each group.

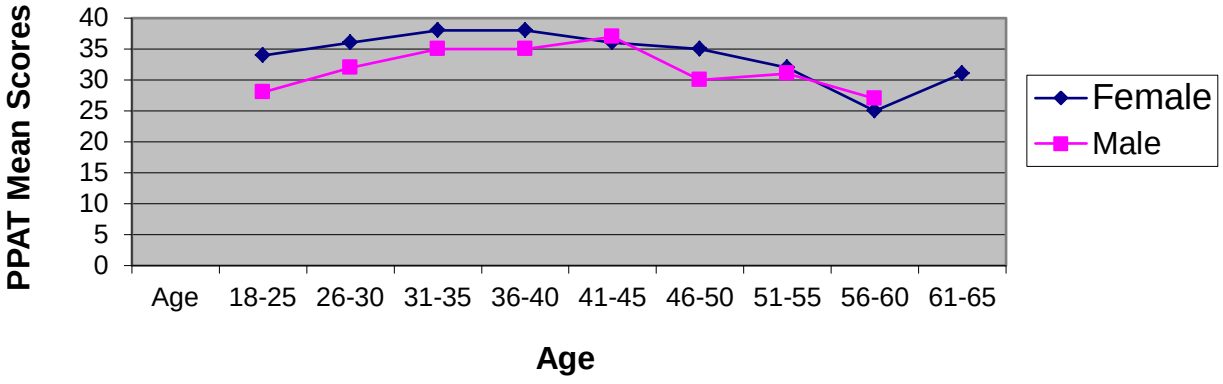

Figure 1. PPAT Gender Comparison

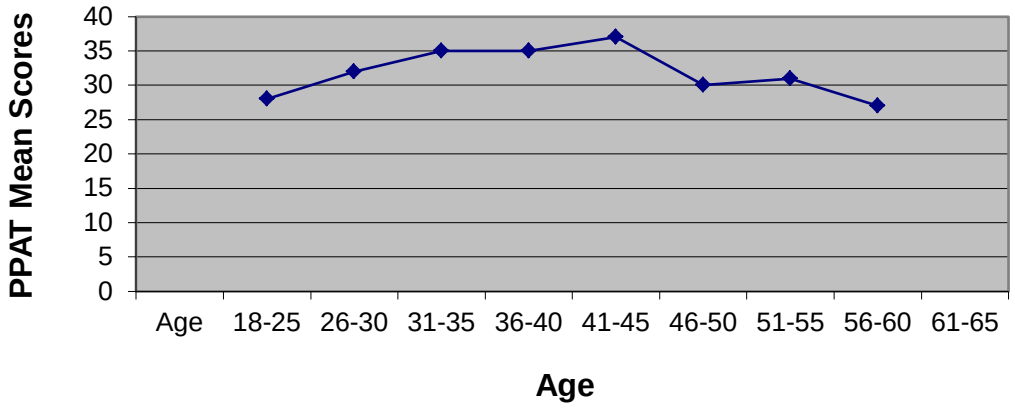

Figure 2. Male PPAT Means by 5 Year Age Categories

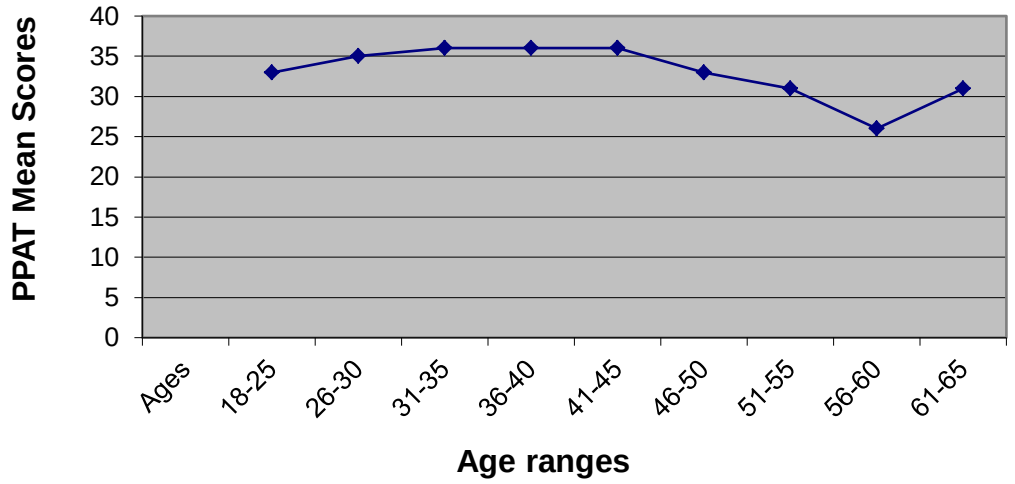

Figure 3. Female PPAT Means by 5 Year Age Categories

Table 3

PPAT Female and Male Norms for Michigan Residents

Female MI Resident Norms				
Age	Mean	N	+- 1SD	SE
18-49	33	56	26-40	.62
50-62	31	14	23-41	1.3
Male MI Resident Norms				
Age	Mean	N	+- 1SD	SE
18-49	31	54	26-36	.44
50-62	27	26	21-33	.80

Discussion

The inter-rater reliability portion of this study included a training session for the student raters and an expanded explanation of scoring procedures. This training manual (including pictures of completed tasks, correct answers according to the scoring rules, and rationale for the correct answers) is available upon request from the first author. The expanded scoring rules allowed better inter-rater reliability and increased the reliability of the test for norm comparison.

The age divisions were split into two age categories, 18 to 49 and 50 to 62 years of age, for both genders, since the mean scores began to decrease in both at about age 49. This could possibly be a result of age-related issues that negatively affect fine motor skills, such as osteoarthritis. The standard of error was higher in both the 50 to 62 year old males (.80) and females (1.3), possibly reflecting the differences in the aging process in different people. This large standard of error for the females may also be due to the small number of female participants in the 50 to 62 year age group (n = 14). Listing the mean ranges for plus or minus one standard deviation allows a clinician to see quickly how a person's score compares to people in his or her age category.

This is the first study that administered the PPAT in isolation of the other three subtests of the PPT. Although using just this one subtest for a general functional capacity test had been reported informally, it was not found in the literature. This study was a preliminary step in the psychometric process of determining whether or not this subtest can be used in isolation as a fairly short yet comprehensive clinical assessment of a person's complex bilateral finger dexterity. The researchers decided to develop general norms using three trials as the first step with the PPAT and the IRR study. The mean of three trials of this subtest was used to insure a more accurate and reliable score (Yancosek & Howell, 2009) since it was not administered with the other subtests of the PPT. The isolated subtest was chosen based on the complexity (Strauss et al., 2006) of the PPAT, and seemed the most likely brief clinical assessment that would correlate with the daily living, vocational, and avocational tasks that people do every day.

Limitations in this study include the small and geographically limited convenience sample with participants that had various types of jobs in the two employment sectors; the research participants did not all have jobs that specifically required finger dexterity. Some of the car factory workers did jobs that required finger dexterity, but others were office workers or assigned to work that required more lifting than dexterity. The health care employees had various jobs that included health care professionals, housekeeping, food service, and office workers. Tiffin and Asher (1948) focused their norms on specific categories of workers who needed to use finger dexterity in their specific job categories, but the researchers for this study felt that more general norms for just the one complex task (PPAT) could be more useful for OTs for comparisons when finger dexterity is needed. Further research is necessary to establish more representative norms through a larger multi-site study. In addition, construct validity studies are needed to validate the PPAT as an effective test of

fine motor/finger dexterity. Research designs should include both the PPAT used alone using the mean of three trials and as a design comparing the PPAT to occupation-centered tasks that incorporate finger dexterity.

Conclusion

This is the first step in establishing the PPAT as an evidence-based assessment that OTs can use when a brief clinical assessment is needed to determine limitations or potential for completing finger dexterity tasks. Good psychometrics and norms would allow clinicians more options in choosing the most appropriate assessment to determine how they can facilitate clients' occupational performance to be more effective and satisfactory, following *The Occupational Therapy Code of Ethics and Ethics Standards* (AOTA, 2010). If the PPAT is an assessment that we choose to continue to use, we, as OTs, have the responsibility to conduct the needed research to support solid evidence-based practice. We cannot wait for the psychologists to conduct the studies to validate and norm the assessment for our use.

DOI: 10.15453/2168-6408.1178

References

Agnew, J., Bolla-Wilson, K., Kawas, C. H., Bleecker, M. (1988). Purdue Pegboard age and sex norms for people 40 years old and older. *Developmental Neuropsychology*, *4*(1), 29-35. http://dx.doi.org/10.1080/87565648809540388

American Occupational Therapy Association. (2010). Occupational therapy code of ethics and ethics standards (2010). *American Journal of Occupational Therapy*, *64*, S17-S26. http://dx.doi.org/10.5014/ajot.2010.64S17

Amirjani, N., Ashworth, N. L., Olson, J. L., Morhart, M., & Chan, K. M. (2011). Validity and reliability of the Purdue Pegboard Test in carpal tunnel syndrome. *Muscle and Nerve*, *43*(2), 171-177. http://dx.doi.org/10.1002/mus.21856

Buddenberg, L. A., & Davis, C. (2000). Test-retest reliability of the Purdue Pegboard Test. *American Journal of Occupational Therapy*, *54*(5), 555-558. http://dx.doi.org/10.5014/ajot.54.5.555

Causby, R., Reed, L., McDonnell, M., & Hillier, S. (2014). Use of objective psychomotor tests in health professionals. *Perceptual & Motor Skills*, *118*(3), 765-804. http://dx.doi.org/10.2466/25.27.PMS.118k27w2

Desrosiers, J., Hébert, R., Bravo, G., & Dutil, E. (1995). The Purdue Pegboard Test: Normative data for people aged 60 and over. *Disability and Rehabilitation*, *17*(5), 217-224. http://dx.doi.org/10.3109/09638289509166638

Doucet, B. M., Woodson, A., & Watford, M. (2014). Centennial Vision—moving toward 2017: Progress in rehabilitation intervention effectiveness research. *American Journal of Occupational Therapy*, *68*, e124-e148. http://dx.doi.org/10.5014/ajot.2014.011874

Edwards, M., Feightner, J., & Goldsmith, C. H. (1995). Inter-rater reliability of assessments administered by individuals with and without a background in health care. *Occupational Therapy Journal of Research*, *15*(2), 103-110. http://dx.doi.org/10.1177/153944929501500203

Fisher, A. G. (2013). Occupation-centred, occupation-based, occupation-focused: Same, same or different? *Scandinavian Journal of Occupational Therapy*, *21*(S1), 96-107. http://dx.doi.org/10.3109/11038128.2014.952912

Fleishman, E. A., & Ellison, G. D. (1962). A factor analysis of fine manipulative tests. *Journal of Applied Psychology*, *46*(2), 96-105. http://dx.doi.org/10.1037/h0038499

Gallus, J., & Mathiowetz, V. (2003). Test-retest reliability of the Purdue Pegboard for persons with multiple sclerosis. *American Journal of Occupational Therapy*, *57*, 108-111. http://dx.doi.org/10.5014/ajot.57.1.108

Lee, P., Liu, C.-H., Fan, C.-W., Lu, C.-P., Lu, W.-S., & Hsieh, C.-L. (2013). The test-retest reliability and the minimal detectable change of the Purdue Pegboard Test in schizophrenia. *Journal of the Formosan Medical Association*, *112*(6), 332-337. http://dx.doi.org/10.1016/j.jfma.2012.02.023

Lindstrom-Hazel, D., Kratt, A., & Bix, L. (2009). Interrater reliability of students using hand and pinch dynamometers. *American Journal of Occupational Therapy*, *63*, 193–197. http://dx.doi.org/10.5014/ajot.63.2.193

Mathiowetz, V., Rogers, S. L., Dowe-Keval, M., Donahoe, L., & Rennells, C. (1986). The Purdue Pegboard: Norms for 14- to 19-year-olds. *American Journal of Occupational Therapy*, *40*(3), 174-179. http://dx.doi.org/10.5014/ajot.40.3.174

National Institutes of Health. (2012). *NIH Toolbox Technical Manual*. Retrieved from http://www.nihtoolbox.org/HowDoI/TechnicalManual/Technical%20Manual%20sections/Toolbox%209-Hole%20Pegboard%20Dexterity%20Test%20Technical%20Manual.pdf

Ottenbacher, K. J., Tickle-Degnen, L., & Hasselkus, B. R. (2002). Therapists awake! The challenge of evidence-based occupational therapy. *American Journal of Occupational Therapy*, *56*(3), 247-249. http://dx.doi.org/10.5014/ajot.56.3.247

Phillips, J. M., Lindstrom-Hazel, D. K., Harrow Swantek, S., & Courtnay Catalano, E. (2013). Grip and pinch strength norms for Michigan workers. *The Open Journal of Occupational Therapy*, *1*(3). http://dx.doi.org/10.15453/2168-6408.1038

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*(2), 420-428. http://dx.doi.org/10.1037/0033-2909.86.2.420

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary* (3rd ed.). New York, NY: Oxford University Press.

Tiffin, J. (1968). *Purdue Pegboard examiner manual*. Chicago, IL: Science Research Associates.

Tiffin, J., & Asher, E. J. (1948). The Purdue Pegboard: Norms and studies of reliability and validity. *Journal of Applied Psychology*, *32*(3), 234-247. http://dx.doi.org/10.1037/h0061266

Yancosek, K. E., & Howell, D. (2009). A narrative review of dexterity assessments. *Journal of Hand Therapy*, *22*(3), 258-270. http://dx.doi.org/10.1016/j.jht.2008.11.004