Published in final edited form as:

Am J Occup Ther. 2012 ; 66(3): 330–337. doi:10.5014/ajot.2012.000935.

Score Reliability and Construct Validity of the Flinn Performance Screening Tool for Adults With Symptoms of Carpal Tunnel Syndrome

Sharon R. Flinn, PhD, OTR/L, CHT, CVE [Assistant Professor],

Division of Occupational Therapy, Ohio State University, 406 Atwell Hall, 453 West 10th Avenue, Columbus, OH 43210; flinn.39@osu.edu

William S. Pease, MD [Professor], and

Ohio State University, Department of Physical Medicine and Rehabilitation, Columbus, OH.

Miriam L. Freimer, MD [Associate Clinical Professor]

Ohio State University, Department of Neurology, Columbus, OH.

Abstract

OBJECTIVE—We investigated the psychometric properties of the Flinn Performance Screening Tool (FPST) for people referred with symptoms of carpal tunnel syndrome (CTS).

METHOD—An occupational therapist collected data from 46 participants who completed the Functional Status Scale (FSS) and FPST after the participants' nerve conduction velocity study to test convergent and contrasted-group validity.

RESULTS—Seventy-four percent of the participants had abnormal nerve conduction studies. Cronbach's α coefficients for subscale and total scores of the FPST ranged from .96 to .98. Intrarater reliability for six shared items of the FSS and the FPST was supported by high agreement (71%) and a fair κ statistic (.36). Strong to moderate positive relationships were found between the FSS and FPST scores. Functional status differed significantly among severe, mild, and negative CTS severity groups.

CONCLUSION—The FPST shows adequate psychometric properties as a client-centered screening tool for occupational performance of people referred for symptoms of CTS.

Keywords

activities of daily living; carpal tunnel syndrome; diagnostic techniques and procedures; psychometrics; reproducibility of results

Carpal tunnel syndrome (CTS) is the most common entrapment neuropathy of the upper extremity (Chan et al., 2007). Compression of the median nerve in the carpal canal of the wrist results in reduced hand function for sensibility, dexterity, pinch, and grip (Amadio, Silverstein, Ilstrup, Schleck, & Jensen, 1996; Atroshi, Johnsson, & Sprinchorn, 1998; Jerosch-Herold, Leite, & Song, 2006). More important, people with CTS report many difficulties in performing everyday activities such as sleeping, buttoning clothes, managing jewelry, opening jars and packages, getting up and down from the floor, and using a computer (Flinn, Pease, & Freimer, 2010). The estimated cost of CTS to society exceeds \$2 billion annually (Palmer & Hanrahan, 1995); CTS affects 1.8 per 1,000 population per year (Bongers, Schellevis, van den Bosch, & van der Zee, 2007).

Diagnosis of CTS depends on the client's description of symptoms and functional limitations. In many circumstances, however, clients have difficulty conveying the specifics

of desired occupations that are challenging for them. Thus, a validated client-centered screening tool that accurately and efficiently identifies functional deficits for the occupational therapy evaluation is advantageous to occupational therapy professionals. To date, research comparing functional status as reported by people with CTS using client-centered card sorts versus the survey formats of standardized questionnaires has been limited.

The Flinn Performance Screening Tool (FPST) shows promise as a comprehensive, client-reported measure of functional status for people with CTS. As a card sort, the FPST is based on the areas of occupation described in the Occupational Therapy Practice Framework: Domain and Process (2nd ed.; American Occupational Therapy Association [AOTA], 2008). The FPST uses 301 photographs and 24 drawings representing seven areas of occupation, 10 categories of activities of daily living (ADLs), 12 categories of instrumental activities of daily living (IADLs), rest, education, work, leisure, and social participation. The client completes the FPST in preparation for the occupational therapy evaluation to create a personal profile of important functions that are difficult to perform. Data collected from the FPST are the basis for goal setting and occupational therapy treatment. The information enables occupational therapy professionals to focus on interventions that address problematic occupations, rather than CTS impairment, and to engage the client with CTS in personally relevant, motivating, and health-promoting activities.

Therefore, our aim was to investigate the psychometric properties of the FPST for people referred with symptoms of CTS by examining the internal consistency, intrarater reliability, and convergent validity of the FPST and evaluating contrasted-group validity for people with different severity levels of CTS.

Method

Research Design

The research design was a methodological study to examine the psychometric properties of the FPST. We examined the internal consistency reliability of the FPST and the intrarater reliability between common items of the FPST and the Functional Status Scale (FSS) of the Boston Carpal Tunnel Questionnaire (Levine et al., 1993). *Score reliability* is defined as the consistency and stability of scores obtained from a specific test for a particular group of people (Thompson, 2003). One type, *internal consistency reliability*, is the extent to which cohesiveness or interrelatedness exists among items in a test (Isaac & Michael, 1997). For a measure with multiple subtests that represent different areas of occupation, such as the FPST, internal consistency is especially important because it reflects the degree to which the item content is analogous (Lemke & Wiersma, 1976). A second type, *intrarater reliability*, is established when a rater completes the same assessment on two or more occasions. When the unit of measurement is a categorical variable, reliability is assessed using a measure of agreement (Portney & Watkins, 1993).

In addition, we investigated convergent validity between the FPST and FSS and contrasted-group validity between the FPST scores of people with different severity levels of CTS. *Construct validity* refers to the extent to which scores on a test measure the hypothetical trait, or construct, the test is intended to measure (Anastasi & Urbina, 1997). This type of validity addresses whether the test effectively measures skills comparable to those sampled by another standardized test with similar constructs (Su, Chen, Tsai, Tsai, & Su, 2007). In this study, all the items measured on the FSS were collected on the FPST. *Convergent validity* is one type of construct validity that addresses whether the test effectively measures skills comparable to those sampled by other standardized tests with similar constructs but different methods of testing (Allen & Yen, 1979). In the current study, we compared ratings

from the FPST with survey responses from the FSS. *Contrasted-group validity*, a second type of construct validity, determines whether a test is able to distinguish among groups that theory claims ought to be discernible (Su et al., 2007). In this study, we expected the magnitude of functional limitations to be different among people with varying degrees of CTS severity.

Instruments

We used three measures to evaluate the functional status and dominance of people with CTS: (1) the FSS, (2) the FPST, and (3) the Waterloo Handedness Questionnaire (WHQ; Bryden, 1977).

Functional Status Scale of the Boston Carpal Tunnel Questionnaire—The FSS of the Boston Carpal Tunnel Questionnaire is a validated, widely accepted, disease-specific measure of ADL performance for people with CTS (Levine et al., 1993). It has been reported to be reproducible, internally consistent, and responsive to clinical change after carpal tunnel surgery (Bessette et al., 1997; Katz et al., 1995, 1998; Levine et al., 1993; You, Simmons, Freivalds, Kothari, & Naidu, 1999). Content validity of the FSS was established by a panel of hand surgeons, rheumatologists, and patients on the basis of the needs of a broad range of CTS clients, age groups, and occupations. Eight common daily activities were used to measure functional status: writing, buttoning clothes, holding a book while reading, gripping the telephone, opening jars, doing household chores, carrying grocery bags, and bathing. The FSS has high test–retest reliability ($r = .93$) and internal consistency ($\alpha = .91$; Levine et al., 1993). Moderate correlations with grip and pinch strength and fair to poor correlation with two-point discrimination and Semmes-Weinstein monofilament testing (Spearman coefficient $r_s = .12$ to $-.42$) have been reported (Katz, Gelberman, Wright, Lew, & Liang 1994; Levine et al., 1993).

FSS items are rated on difficulty levels using a 5-point Likert scale with 0 indicating *no difficulty* and 5 indicating *very severe difficulty or inability to perform*. The scores on the FSS ranged from 8 to 32, with higher scores indicating greater limitations in functional performance.

Flinn Performance Screening Tool—The FPST evaluates functional status on the basis of client responses to 301 photographs depicting ADLs, IADLs, rest, education, volunteerism, leisure, and social participation. Twenty-four drawings are used for work but were not included in this study. Personal hygiene, one example of an ADL category, is represented by photos of people washing hair, caring for toenails, and applying deodorant. Meal preparation, one example of an IADL category, is represented by photos of people using a microwave, opening and closing a jar, and washing silverware. Rest, education, volunteerism, leisure, and social participation are included in the IADL section.

Sampling validity of the FPST was established using the categories of occupation from the *Framework* to represent all of the domains of function. Face validity was established from the input of 200 orthopedic and rheumatology clients. Test–retest reliability of the FPST was reported for healthy occupational therapy students, with 92% of the items meeting a Kendall's τ coefficient greater than .80 and 97% agreement on two test conditions (Flinn, Ventura, & Goodman, 1990). Acceptable internal consistency ($\alpha = .89$) and interrater agreement were found between veterans with severe mental illness and their occupational therapy personnel, $r(34) = .483$, $p = .002$ (Flinn, 2004).

The FPST has been used to identify the functional limitations reported by people with upper-extremity orthopedic diagnoses, including CTS, who were referred to an industrial rehabilitation program (Flinn-Wagner, Mladonicky, & Goodman, 1990). No significant

differences were found in the number of ADL issues for the return to work (RTW) and non-return to work (NRTW) groups. However, the participants in both groups had a large number of deficits at admission (RTW mean = 41, range = 0–100; NRTW mean = 37, range = 8–72). In a second study, clients with carpal tunnel releases were asked to identify important but difficult ADLs and IADLs using the FPST (Flinn, Ventura, & Goodman, 2005). Clients reported limitations in all areas of occupation except spirituality. Seventy-five percent of all items were identified by at least 1 participant as important but poorly performed. A total of 1,420 deficits were recorded among the 29 participants, with an average of 48.96 items per subject (standard deviation [*SD*] = 40.49, range = 0–175). These findings suggest that many valued life activities are not reported in standard outcome measures, which can lead practitioners to underestimate the upper-limb disability caused by CTS.

A dichotomous rating scale is used for the FPST in which 1 indicates a functional problem and 0 indicates no functional problem or a lack of importance to the client. The scores on the FPST range from 0–144 for the ADL subscale, 0–157 for the IADL subscale, and 0–301 for the total scale. The score for the FPST is determined by the total number of items selected. Therefore, the score on the FPST represents items the client describes as important but problematic, and higher scores indicate greater extent of disability.

Waterloo Handedness Questionnaire—In addition to the two measures of functional status, we used the WHQ to measure hand preference (Bryden, 1977). The WHQ has established intersubject (.69) and interitem (.88) reliability (Fess, 1997). Concurrent criterion-referenced validity was established (Brown, Roy, Rohr, Snider, & Bryden, 2004) for healthy people between scores on the WHQ and the following performance-based indicators of hand preference: Wathand Box ($r = .881$), Grooved Pegboard ($r = .750$), Annett Pegboard ($r = .614$), grip strength ($r = .456$), and finger tapping ($r = .405$). In a second validation study with healthy participants (Bryden & Roy, 2005), the WHQ correlated with a multidimensional observational test of handedness using the Wathand Cabinet Test ($r = .795$, $p = .01$). In general, the preferred hand was faster than the nonpreferred hand among healthy people for the place test on the Grooved Pegboard, $F(1, 149) = 587.7$, $p < .001$ (Bryden, Roy, & Spence, 2007).

Clients complete a 19-item questionnaire using a 5-point Likert scale: *always use the left hand* (–2), *usually use the left hand* (–1), *use both hands equally often* (0), *usually use the right hand* (1), and *always use the right hand* (2). Scores used to categorize participants are –38 to –20 for left handed, –19 to 19 for ambidextrous, and 20 to 38 for right handed (Ashworth, Ciorciari, & Stough, 2008).

Participant Selection

Participants were referred to one of two electromyography (EMG) clinics at a large medical center in central Ohio for differential diagnosis of CTS. After the testing procedure, the physical medicine and rehabilitation physician or neurologist recruited participants for the study. Because the participants had already been evaluated by their primary care physician or hand surgeon, no clients were excluded, and the inclusion criteria were completion of an EMG for median neuropathy and age of 18–70 yr.

Procedures

The study received ethics clearance from the behavioral and social sciences institutional review board at Ohio State University. Informed consent was obtained from each participant. Either a physiatrist or a neurologist performed electrodiagnostic testing using standardized protocols. Comparative nerve conduction studies were recordings between

median-to-ulnar and median-to-radial tests. Segmental studies were recordings between the digit–palm and palm–wrist segments, which yield higher sensitivity in detecting mild CTS (Sheu, Chiou, Hu, & Chen, 2006). The physician authors of this article (William S. Pease and Miriam L. Freimer) reviewed the nerve conduction reports from their respective departments and assigned one of six severity categories to each hand (Padua et al., 1997). A description of the neurophysiological classification used to assign CTS severity categories is provided in Table 1.

Data Collection

Once we obtained consent, Sharon R. Flinn collected data. Demographic information collected included age, gender, ethnic background, level of education, workers' compensation claims, hand preference, problem hand, duration of hand symptoms (in months), CTS severity, related comorbidities, and body mass index (BMI). We were each blinded to the data collected by each other.

Data Analysis

We used descriptive statistics to summarize the demographic information and severity of CTS symptoms by problematic hand. We calculated Cronbach's α coefficients to evaluate the internal consistency of the FPST and FSS. An intrarater reliability analysis using cross-tabulation and Cohen's κ statistics was selected to provide an unbiased evaluation of consistency among each participant's responses for six common items in the FPST and FSS. Those items with $\geq 70\%$ agreement were considered to have a high level of agreement, whereas items with $<70\%$ agreement were considered to have low agreement (Aday, 1989; Fleiss, 1981). Item κ values of .81–1.00 were interpreted as almost perfect agreement, .61–.80 as substantial agreement, .41–.60 as moderate agreement, .21–.40 as fair agreement, .00–.20 as slight agreement, and <0.00 as poor agreement (Landis & Koch, 1977).

We evaluated convergent validity by computing a Pearson product–moment correlation coefficient between the scores of the FSS and the subscale and total scores of the FPST. We evaluated contrasted-group validity using a one-way between-subjects analysis of variance to compare the effect of CTS severity on four functional outcome variables: (1) the FSS, (2) the FPST ADL subscale, (3) the FPST IADL subscale, and (4) the FPST total score. Post hoc differences were analyzed among severe, mild, and negative CTS groups using Fisher's least significant difference (LSD). Statistical analysis was performed with IBM SPSS Statistics 19.0 (IBM, Armonk, NY), and statistical significance was set at $p < .05$.

Results

A sample size of 45 participants was estimated a priori to provide adequate power on the basis of an analysis conducted using G*Power 3.0.8 (Erdfelder, Faul, & Buchner, 1996). With a sample of 45 and an $\alpha = .05$, this design could detect validity coefficients as low as .4 with a calculated power of .8. Complete data were collected on 46 participants. Forty-seven participants consented to the study, but 1 withdrew because of time constraints and was not included in the analyses. Sample characteristics are provided in Table 2.

Adequate score reliability was found for both measures of functional status. Cronbach's α coefficients were .906 for the FSS, .962 for the FPST ADL subscale, .963 for the FPST IADL subscale, and .980 for the FPST total scores. The proportion of agreement between items was high for four of six items. A fifth item, carrying groceries, had high agreement if the photograph represented carrying groceries using a plastic bag with handle grips rather than using a paper bag with no handle grips. The sixth item, opening a jar, had only

moderate agreement. In spite of high agreement between most items, the average κ scores were fair. Table 3 outlines intrarater reliability results by item.

Convergent validity was supported because moderate to strong positive relationships were found between the scores of the FSS and the FPST ADL subscale ($r = .741, n = 42, p < .001$); the FSS and the FPST IADL subscale ($r = .619, n = 41, p < .001$); and the FSS and the FPST total ($r = .725, n = 45, p < .001$).

Contrasted-group validity was supported; significant differences were found in all four functional outcome variables at the $p < .05$ level for the three severity conditions. Functional status differed significantly across the CTS severity levels for the FFS, $F(2, 43) = 3.53, p = .035$; the FPST ADL subscale, $F(2, 39) = 4.20, p = .022$; the FPST IADL subscale, $F(2, 38) = 3.29, p = .048$; and the FPST total, $F(2, 42) = 4.04, p = .025$. Post hoc comparisons using the Fisher's LSD test indicated significant differences for both measures between the severe and mild CTS groups and between the severe and negative CTS groups. However, no significant effects were found on either measure between the mild and negative CTS groups. Post hoc comparisons for the FSS and FPST subscale and total scores by severity group are reported in Table 4.

The demographic characteristics were not significantly different between the CTS and non-CTS groups, with two exceptions: (1) For ethnic background, 73.5% of CTS participants versus 4% of non-CTS participants were White, $\chi^2(3, N = 46) = 13.50, p = .004$, and (2) for body mass index, 67% of CTS participants versus 50% of non-CTS participants were obese, $\chi^2(3, N = 45) = 14.97, p = .002$. Both CTS and non-CTS participants had experienced symptoms for an average of 8 mo. About 50% of the participants complained of back, neck, or shoulder symptoms, and 80% had abnormal BMI. Most important, the non-CTS group had an average of 24 ADL and 17 IADL limitations. This finding suggests that despite the absence of neurophysiological impairment at the wrist, limitations in occupational performance are present and may be an early indicator of pathology. If so, an opportunity could be afforded by occupational therapy personnel to reduce the risk factors associated with CTS through a program of preventative care and lifestyle changes.

Discussion

The findings of this study support the score reliability and construct validity of the FPST for people referred for symptoms of CTS. Internal consistency of the FPST and FSS subscale and total scores was .906–.980, which supports the results of previous studies on the FSS (Levine et al., 1993).

We examined interrater reliability by administering the FSS before the FPST. We used Cohen's κ statistic to analyze the data by comparing the response for a specific functional task on the FSS with the response for the same task on the FPST. The average percentage of agreement among common items was high, whereas the κ coefficients demonstrated only fair intrarater reliability. Writing was the only item with high agreement and a substantially high κ coefficient. High agreement but low κ was observed for buttoning, carrying plastic bags, holding a book, and gripping a telephone. Feinstein and Cicchetti (1990) suggested that the reason for this difference is an imbalance between the chance-corrected observed and expected agreements or the presence of asymmetrical differences in the marginal totals. In reviewing the data, we found that the four items demonstrated this phenomenon of high difficulty for one item or the other. Of interest was high intrarater agreement for endurance activities versus low agreement for strength-related activities such as opening a jar. In addition, agreement differed on the basis of type of grocery bag carried (plastic vs. paper). These findings suggest the advantage of taking a client-centered approach and identifying

the activity demands of each functional task as accurately as possible when measuring functional status.

Findings indicate that the FSS and FPST measured similar constructs with the use of different testing methods. We found convergent validity between the FSS and the FPST and a moderate correlation between the scores of the FSS and the subscale and total scores of the FPST ($r_s = .619-.741$). The weaker relationship we found between the FSS and the IADL subscale of the FPST was expected because the FSS items largely represent ADL areas of occupation.

Contrasted-group validity was supported for both the FSS and the FPST; the mean scores for each measure differed significantly on the basis of CTS severity group membership. However, post hoc analysis found that functional limitations differed significantly when the discrepancies between impairment levels were larger. For example, clients in the severe and negative membership groups showed greater variances than the severe and mild groups when examining functional limitations.

Implications for Occupational Therapy Practice

Increased use of screening tools in occupational therapy can be an efficient and effective method of gathering information for a client's occupational profile and assessment. The use of a client-centered card sort enhances the skills of an occupational therapy practitioner to provide meaningful intervention for clients and to focus the efforts of rehabilitation team members on functional outcomes.

- The results of the study provide evidence of the score reliability and construct validity of the FPST for people with CTS symptoms.
- As a screening tool, the FPST provides specific detail from a comprehensive menu of occupations that can guide the occupational therapy evaluation, support collaborative goal activities, contribute to a custom therapy program for activity modification, and demonstrate measurable functional outcomes.
- In addition, the photographs from the FPST provide advantages that aid cognitive processes, generate client narratives, and save time in the evaluation process (Anzi, 1984; Baum & Edwards, 1993; Berg & LaVesser, 2006; Matheson, 2004).
- Occupational therapists should consider using client-centered screening tools that use card sorts to build a personalized profile with the client for function-driven occupational therapy services.

Limitations and Future Research

Limitations of the study included a small sample size and disproportionate membership in the extreme, severe, and minimal severity groups. Strengths of the study design were the use of nerve conduction studies as the gold standard to classify participants by CTS severity and those who were potential candidates for conservative CTS treatment.

Future studies that examine the characteristics of larger samples of people with milder or negative CTS are recommended, especially studies on functional challenges by severity group, risk factors associated with nonwork functional limitations, and the influence of multiple orthopedic comorbidities and high incidences of ambidexterity on the pathogenesis of CTS.

Acknowledgments

We thank Jane Case-Smith, Susan White, Christopher Taylor, and Ranata Topp for their editorial and statistical assistance.

References

- Aday, L. Designing and conducting health surveys: A comprehensive guide. Jossey-Bass; San Francisco: 1989.
- Allen, M.; Yen, W. Introduction to measurement theory. Brooks/Cole; Monterey, CA: 1979.
- Amadio P, Silverstein M, Ilstrup D, Schleck C, Jensen L. Outcome assessment for carpal tunnel surgery: The relative responsiveness of generic, arthritis-specific, disease-specific, and physical examination measures. *Journal of Hand Surgery*. 1996; 21A:338–346. [PubMed: 8724457]
- American Occupational Therapy Association. Occupational therapy practice framework: Domain and process (2nd ed.). *American Journal of Occupational Therapy*. 2008; 62:625–683. <http://dx.doi.org/10.5014/ajot.62.6.625>. [PubMed: 19024744]
- Anastasi, A.; Urbina, S. Psychological testing. 7th ed. Prentice Hall; Upper Saddle River, NJ: 1997.
- Anzi, D. Loma Linda Activity Sort. Work Evaluation Systems Technology; Fort Bragg, CA: 1984.
- Ashworth J, Ciorciari J, Stough C. Psychophysiological correlates of dissociation, handedness, and hemispheric lateralization. *Journal of Nervous and Mental Disease*. 2008; 196:411–416. <http://dx.doi.org/10.1097/NMD.0b013e31816ff384>. [PubMed: 18477884]
- Atroshi I, Johnsson R, Sprinchorn A. Self-administered outcome instrument in carpal tunnel syndrome. *Acta Orthopaedica Scandinavica*. 1998; 69:82–88. <http://dx.doi.org/10.3109/17453679809002363>. [PubMed: 9524525]
- Baum, M.; Edwards, D. Activity Card Sort. Washington University School of Medicine, Occupational Therapy Program; St. Louis, MO: 1993.
- Berg C, LaVesser P. The Preschool Activity Card Sort. *OTJR: Occupation, Participation and Health*. 2006; 26:143–151.
- Bessette L, Keller R, Liang M, Simmons B, Fossel A, Katz J. Patients' preferences and their relationship with satisfaction following carpal tunnel release. *Journal of Hand Surgery*. 1997; 22A: 613–620. [PubMed: 9260615]
- Bongers F, Schellevis F, van den Bosch W, van der Zee J. Carpal tunnel syndrome in general practice (1987 and 2001): Incidence and the role of occupational and non-occupational factors. *British Journal of General Practice*. 2007; 1:36–39. [PubMed: 17244422]
- Brown S, Roy E, Rohr L, Snider B, Bryden P. Preference and performance measures of handedness. *Brain and Cognition*. 2004; 55:283–285. <http://dx.doi.org/10.1016/j.bandc.2004.02.010>. [PubMed: 15177796]
- Bryden M. Measuring handedness with questionnaires. *Neuropsychologia*. 1977; 15:617–624. [http://dx.doi.org/10.1016/0028-3932\(77\)90067-7](http://dx.doi.org/10.1016/0028-3932(77)90067-7). [PubMed: 896019]
- Bryden P, Roy E. A new method of administering the Grooved Pegboard Test: Performance as a function of handedness and sex. *Brain and Cognition*. 2005; 58:258–268. doi:10.1016/j.bandc.2004.12.004. [PubMed: 15963376]
- Bryden P, Roy E, Spence J. An observational method of assessing handedness in children and adults. *Developmental Neuropsychology*. 2007; 32:825–846. doi: 10.1080/87565640701539667. [PubMed: 17956184]
- Chan L, Turner J, Comstock B, Levenson L, Hollingsworth W, Heagerty P, et al. The relationship between electrodiagnostic findings and patient symptoms and function in carpal tunnel syndrome. *Archives of Physical Medicine and Rehabilitation*. 2007; 88:19–24. <http://dx.doi.org/10.1016/j.apmr.2006.10.013>. [PubMed: 17207670]
- Erdfelder E, Faul F, Buchner A. GPOWER: A general power analysis program. *Behavior Research Methods, Instruments, & Computers*. 1996; 28:1–11. <http://dx.doi.org/10.3758/BF03203630>.
- Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. *Journal of Clinical Epidemiology*. 1990; 43:543–549. [http://dx.doi.org/10.1016/0895-4356\(90\)90158-L](http://dx.doi.org/10.1016/0895-4356(90)90158-L). [PubMed: 2348207]

- Fess EE. Defining handedness: A critical variable. *Journal of Hand Surgery*. 1997; 10:314.
- Fleiss, J. Statistical methods for rates and proportions. Wiley; New York: 1981.
- Flinn S. Reliability and validity of the Recovery Assessment Scale for consumers with severe mental illness living in group home settings. *Dissertation Abstracts International*. 2004; 65(11):4115. (UMI No. 3152236).
- Flinn, S.; Pease, W.; Freimer, M. Characteristics of physical and functional variables by CTS severity. Poster presented at the annual conference of the International Federation of Societies for Hand Therapy; Orlando, FL. Jun. 2010
- Flinn, S.; Ventura, D.; Goodman, G. Reliability of the Flinn Performance Screening Tool, a client centered instrument to identify functional deficits. 1990. Unpublished manuscript
- Flinn, S.; Ventura, D.; Goodman, G. ADL issues commonly reported by individuals with carpal tunnel releases: A descriptive study. Poster presented at the annual conference of the American Society of Hand Therapy; San Antonio, TX. Sep. 2005
- Flinn-Wagner S, Mladonicky A, Goodman G. Characteristics of upper extremity injured individuals making a successful transition to work. *Journal of Hand Therapy*. 1990; 2:51–55.
- Isaac, S.; Michael, W. Handbook in research and evaluation: A collection of principles, methods, and strategies useful in the planning, design, and evaluation of studies in education and behavioral sciences. 3rd ed. EdITS; San Diego, CA: 1997.
- Jerosch-Herold C, Leite J, Song F. A systematic review of outcomes assessed in randomized controlled trials of surgical intervention for carpal tunnel syndrome using the International Classification of Functioning, Disability, and Health (ICF) as a reference tool. *BMC Musculoskeletal Disorders*. 2006; 7:96. <http://dx.doi.org/10.1186/1471-2474-7-96>. [PubMed: 17147807]
- Katz J, Fossel K, Simmons B, Swartz R, Fossel A, Koris M. Symptoms, functional status, and neuromuscular impairment following carpal tunnel release. *Journal of Hand Surgery*. 1995; 20A: 549–555. [PubMed: 7594277]
- Katz J, Gelberman R, Wright E, Lew R, Liang M. Responsiveness of self-reported and objective measures of disease severity in carpal tunnel syndrome. *Medical Care*. 1994; 32:1127–1133. <http://dx.doi.org/10.1097/00005650-199411000-00005>. [PubMed: 7967853]
- Katz J, Keller R, Simmons B, Rogers W, Bessette L, Foy S, et al. Maine Carpal Tunnel Study: Outcomes of operative and non-operative therapy for carpal tunnel syndrome in a community-based cohort. *Journal of Hand Surgery*. 1998; 23A:697–710. [PubMed: 9708386]
- Landis J, Koch G. The measurement of observer agreement for categorical data. *Biometrics*. 1977; 33:159–174. <http://dx.doi.org/10.2307/2529310>. [PubMed: 843571]
- Lemke, E.; Wiersma, W. Principles of psychological measurement. Rand McNally; Chicago: 1976.
- Levine D, Simmons B, Koris M, Daltroy L, Hohl G, Fossel A, et al. A self-administered questionnaire for the assessment of severity of symptoms and functional status in carpal tunnel syndrome. *Journal of Bone and Joint Surgery*. 1993; 75:1585–1592. [PubMed: 8245050]
- Matheson LN. History, design characteristics, and uses of the pictorial activity and ask sorts. *Journal of Occupational Rehabilitation*. 2004; 14:175–195. <http://dx.doi.org/10.1023/B:JOOR.0000022760.71820.91>. [PubMed: 15156777]
- Padua L, LoMonaco M, Gregori B, Valente E, Padua R, Tonali P. Neurophysiological classification and sensitivity in 500 carpal tunnel hands. *Acta Neurologica Scandinavica*. 1997; 96:211–217. [PubMed: 9325471]
- Palmer D, Hanrahan L. Social and economic costs of carpal tunnel surgery. *Instructional Course Lectures*. 1995; 44:167–172. [PubMed: 7797856]
- Portney, L.; Watkins, M. Foundations of clinical research: Applications to practice. Appleton & Lange; Norwalk, CT: 1993.
- Sheu J, Chiou H, Hu C, Chen W. Segmental study of the median nerve versus comparative tests in the diagnosis of mild carpal tunnel syndrome. *Clinical Neurophysiology*. 2006; 117:1249–1255. <http://dx.doi.org/10.1016/j.clinph.2006.02.004>. [PubMed: 16600674]
- Su C, Chen W, Tsai P, Tsai C, Su W. Psychometric properties of the Loewenstein Occupational Therapy Cognitive Assessment–Second Edition in Taiwanese persons with schizophrenia.

American Journal of Occupational Therapy. 2007; 61:108–118. <http://dx.doi.org/10.5014/ajot.61.1.108>. [PubMed: 17302112]

Thompson, B. Score reliability: Contemporary thinking on reliability issues. Sage; Thousand Oaks, CA: 2003.

You H, Simmons Z, Freivalds A, Kothari M, Naidu S. Relationship between clinical symptom severity scales and nerve conduction measures in carpal tunnel syndrome. Muscle and Nerve. 1999; 22:497–501. doi: 10.1002/(SICI)1097-4598(199904)22:4<497::AID-MUS11>3.0.CO;2-T. [PubMed: 10204785]

Table 1

Carpal Tunnel Syndrome Severity Categories

Category	Description
Extreme	Absence of median thenar motor and sensory responses
Severe	Absence of median sensory response and abnormal distal motor latencies
Moderate	Slowing of median sensory response and abnormal distal motor latencies
Mild	Slowing of median sensory response and normal distal motor latencies
Minimal	Normal median sensory response but abnormal responses in segmental tests
Negative	Normal findings in comparative or segmental tests

Table 2Sample Characteristics (*N* = 46)

Variable	<i>M</i> (<i>SD</i>)	<i>n</i>	%
Age (yr)	47 (10.6)		
Gender (female)		34	73.9
Ethnicity			
White		30	65.2
Black		10	21.7
Other		6	13.0
Education			
High school		16	34.8
Some college		9	19.6
College graduate		21	45.7
Workers' compensation claim		10	21.7
Hand preference			
Right		34	73.9
Left		3	6.5
Equal		9	19.6
Problem hand			
Right		25	54.3
Left		12	26.1
Both		9	19.6
Length of symptoms, mo (median IQR)	8 (3–24)		
Other medical history			
Back symptoms		24	52.2
Neck symptoms		23	50.0
Shoulder symptoms		20	43.5
Diabetes		8	17.4
Inflammatory arthritis		13	28.3
Menopause		12	26.1
Wrist or hand trauma		11	23.9
Body mass index			
Normal		9	19.6
Overweight		12	26.1
Obese		23	50.0
Extremely obese		1	2.2
Missing data		1	2.2
Severity group			
Severe		2	4.3
Mild		31	67.4
Negative		13	28.3

Note. IQR = interquartile range; *M* = mean; *SD* = standard deviation.

Table 3Intrater Reliability Results, by Item ($N = 46$)

Item	Intrater Reliability				
	Agreement ^a	κ Statistic			
		Value ^b	SE	95% CI	<i>p</i>
Writing	.89	.71	.12	.47–.95	<.001
Buttoning	.72	.41	.13	.15–.67	.002
Carrying groceries (plastic bag)	.72	.37	.13	.11–.63	.005
Holding a book	.72	.27	.15	–.03–.57	.045
Gripping telephone	.70	.31	.14	.03–.59	.024
Carrying groceries (paper bag)	.65	.26	.14	–.02–.54	.073
Opening a jar	.57	.21	.08	.05–.37	.021
Total item average	.71	.36			

Note. CI = confidence interval; SE = standard error.

^aHigh = 70% agreement; low = <70% agreement.

^bValues are substantial, .61–.81; moderate, .41–.60; or fair, .21–.40.

Table 4

Post Hoc Comparisons for FSS and FPST Subscale and Total Scores, by Severity Group

					95% CI	
Assessment and Severity Group	Mean Difference	SE	p	Lower Bound	Upper Bound	
FSS						
Severe vs. mild	13.71	5.10	.010	3.43	23.99	
Severe vs. negative	13.31	5.31	.016	2.61	24.01	
Mild vs. negative	−0.40	2.31	.863	−5.06	4.25	
FPST ADL subscale						
Severe vs. mild	38.31	13.22	.006	11.57	65.06	
Severe vs. negative	36.55	13.90	.012	8.42	64.67	
Mild vs. negative	−1.765	6.40	.784	−14.72	11.19	
FPST IADL subscale						
Severe vs. mild	28.54	13.68	.044	0.83	56.25	
Severe vs. negative	36.17	14.27	.015	7.29	65.05	
Mild vs. negative	7.63	6.48	.246	−5.49	20.75	
FPST total						
Severe vs. mild	70.10	26.39	.011	16.84	123.36	
Severe vs. negative	77.81	27.45	.007	22.41	133.20	
Mild vs. negative	7.71	12.00	.524	−16.51	31.93	

Note. ADL = activity of daily living; CI = confidence interval; FPST = Flinn Performance Screening Tool; FSS = Functional Status Scale; IADL = instrumental activity of daily living; SE = standard error.