2014

A study of inter-instrument reliability and concurrent validity of Baseline digital pinch gauge and the B&L pinch gauge

Kailee Miller

A Study of Inter-instrument Reliability and Concurrent Validity of the Baseline® Digital Pinch Gauge and the B&L Pinch Gauge

Kailee Miller

Research Advisor: Julie Thomas, Ph.D., OTR/L, FAOTA

Occupational Therapy Doctorate Program

Department of Rehabilitation Sciences

The University of Toledo

May 2014

This scholarly project reflects individualized, original research conducted in partial fulfillment of the requirements for the Occupational Therapy Doctorate Program, The University of Toledo.

Abstract

Introduction: Occupational therapists commonly use pinch gauges to measure pinch strength in therapy. No comparison research has been done on the Baseline digital pinch gauge. The research question we are testing is, what is the degree of inter-instrument reliability and concurrent validity between the Baseline digital pinch gauge and the B&L hydraulic pinch gauge?

Methods: A total of 102 subjects between the ages of 18 and 50 were recruited for the study. All participants were self reported to be healthy and had no history of neuromuscular or orthopedic conditions that affected their upper extremity function. This study was part of a larger study which included testing of the Baseline digital dynamometer as well. Researchers randomized the presentation of the two pinch gauges and two dynamometers to control for order effect. We tested three pinch patterns: lateral, three-jaw chuck, and tip pinch.

Results: Analyses indicated that there were no significant order effects for the paired data. Data analysis using a two-way mixed effects model ICC demonstrated significant correlations of 0.95 for three-jaw chuck, 0.97 for lateral pinch, and 0.86 for tip pinch. Comparison of means between instruments using t-tests for paired measures found statistically significant differences between the B&L and Baseline Digital pinch gauge for all pinch patterns ($p < 0.05$). Effect sizes were small to moderate.

Discussion: Significant findings of the t-test demonstrate that concurrent validity did not occur between the B&L and Baseline pinch gauge. The results of this study indicate that the Baseline Digital pinch gauge is not interchangeable with the B&L hydraulic pinch gauge.

A Study of Inter-instrument Reliability and Concurrent Validity of the Baseline® Digital Pinch Gauge

Occupational therapists commonly use pinch gauges to measure pinch strength in therapy. Pinch strength is typically taken as a way to compare the individual's strength to normative standards of individuals of the same sex and age (Kellor et al., 1971; Mathiowetz, Vizenor, & Melander, 2000; Puh, 2010). The B&L pinch gauge, a type of hydraulic pinch gauge, is commonly used in clinics, but digital pinch gauges are now available to clinicians. A comparison of the Baseline digital pinch gauge and the B&L pinch gauge was done in the present study to determine if there are differences in pinch strength measurements between the two instruments. If the Baseline digital pinch gauge has inter-instrument reliability and concurrent validity with the B&L pinch gauge then occupational therapists can confidently use the B&L normative data. In order to understand the present study, prior research on pinch strength assessments and instruments including use of various pinch gauges will be reviewed as well as research on the interchangeability of these instruments.

Why is pinch strength important?

One of the most important functions of the hand is the ability to manipulate objects (Ranganthan, Siemionow, Sahgal, & Yue, 2001). When a disability or impairment occurs that affects a person's hand it can impair the individual's ability to perform daily occupations. Impairment of the hand is commonly measured using pinch strength. Pinch strength is taken as a way to compare the individual's strength to normative standards of individuals of the same sex and age (Kellor et al., 1971; Mathiowetz, Vizenor, & Melander, 2000; Puh, 2010). Pinch strength as measured by a pinch gauge is one of the most common assessments used due to being quick, reliable, easy to perform, and easily recordable (Puh, 2010). Pinch strength can be used to show improvements or deteriorations following an intervention (Brorson, Werner, & Thorngren, 1989; Fess, 1986; Mathiowetz et al., 1985; Mathiowetz, Wiemer & Federman, 1986). Pinch strength can also be used as a motivational tool for the client (Brorson et al., 1989; Mathiowetz et al., 2000), an assessment tool for upper limb impairments (Puh, 2010), and a way to create realistic treatment goals (Mathiowetz, Weber, Volland, & Kashman, 1984; Mathiowetz et al., 1985; Mathiowetz et al., 1986).

Smith and Benge (1985) investigated the procedures that occupational therapists use when taking pinch strength measurements. Clinicians across the country use the various normative standards available (Smith & Benge, 1985). Smith and Benge (1985) report that 61.9% of clinicians they surveyed used three trials when measuring pinch strength, 14.3% used two trials, 13.8% used one trial, and 9.5% used a combination of these. Most clinicians tested clients sitting down (80.5%), but some perform the tests without instructing the individual on whether to sit or stand (Smith & Benge, 1985). The author reports inconsistencies in arm positioning for testing, but a majority of the respondents reported having the elbow flexed at 90 degrees and the forearm in neutral (Smith & Benge, 1985). Although the authors did not find one procedure that all clinicians used, a majority that they surveyed performed pinch strength with three trials, the client sitting down and the elbow flexed at 90 degrees while the forearm is in neutral.

Smith & Benge (1985) found inconsistencies in prehension patterns and terminology used with clinicians and researchers. The most common prehension patterns and terminology used were the three-jaw chuck, tip, and lateral pinch (Kellor et al., 1971; Smith & Benge, 1985). Common occupations performed using these prehension patterns are holding a key, tying shoes, fastening buttons, feeding, grooming, putting on earrings, writing, and picking up objects (Ranganathan et al., 2001; Rice, Leonard & Carter, 1998). The most common pinch used in about 60% of occupations of daily living is the three jaw chuck pinch (also known as three point or palmar pinch) (Su, Chein, Cheng & Su, 1995). Of the 195 questionnaires returned in the Smith and Benge (1985) study, 13.6% of respondents called three-jaw chuck a three point pinch, and 8.4% called it a palmar pinch. Other names used for tip pinch are tip to tip and fingertip pinch (Smith & Benge, 1985). Lateral pinch is also commonly called key pinch (Smith and Benge, 1985).

In summary, researchers have found that occupational therapists use inconsistent terms for describing pinch types and inconsistent protocols for measuring pinch strength. Use of inconsistent protocols to measure pinch strength can contribute to measurement error and result in unreliability of results. When comparing a new pinch instrument to a previous pinch instrument, researchers should follow the same protocols and use the same terminology for both instruments.

Reliability and Validity of Pinch Gauges

Validity is defined as the truthfulness of an assessment tool, or the degree to which it measures what it is supposed to measure (Fess, 1986). Comparing a new instrument to a previous instrument that has been found valid provides a measure of the new instrument's concurrent validity. Concurrent validity occurs when the results of the new instrument agrees with the result of the established instrument that has already shown to be valid (Kielhofner, 2006). If a "gold standard" exists, concurrent validity is the best method of choice for evaluating a new instrument

(Kielhofner, 2006). The "gold standard" for pinch gauges is the B&L pinch gauge due to its high reliability and validity. It also has established norms for males and females by age groupings.

Reliability is the ability of an instrument to measure consistently and predictably (Fess, 1986). Reliability is usually expressed in correlation coefficients and standard errors of measurement (SEM) (Fess, 1986). If an instrument has high reliability it gives consistently accurate and predictable outcomes between sessions, referred to as test-retest reliability and between examiners, referred to as inter-rater reliability (Fess, 1986). Research has found that higher test-retest reliability in pinch strength measurements occurs when the mean of three trials is taken instead of one trial only (Mathiowetz et al., 1986). Intra-instrument reliability can be measured by testing the pinch gauges prior to and after the study by determining the relationship between the gauge readout and force applied (usually hanging a known weight from the gauge) (Mathiowetz et al. 2000). When the therapist does not use standardized positioning of the upper extremity during measurements or standardized verbal instructions for administering the test, it can lead to unreliability (Mathiowetz et al., 1984). Use of different instruments for monitoring the same patient can contribute to unreliability (Flood- Joy & Mathiowetz, 1987). Proper calibration of the pinch gauges can contribute to increased reliability (Niebuhr, Marion & Fike, 1994).

There is no one set of standardized protocols and terminology for measuring pinch strength that occupational therapists follow universally. Having dissimilar protocols and terminology can affect the reliability and validity of pinch strength measurements (Smith & Benge, 1985). When testing for reliability between instruments (inter-instrument reliability) the same protocols need to be followed. Upon review of the literature, no concurrent validity research has been reported between the B&L hydraulic pinch gauge and the Baseline digital pinch gauge.

Interchangeability

Normative data needs to be developed for pinch gauges to be useful for patient assessment. Likewise, as new pinch gauges come onto the market, they need to be compared to existing pinch gauges to determine concurrent validity and reliability (Mathiowetz, 2000; Fess, 1986; Hinson, Woodard & Gench, 1990). Kellor et al. (1971) created the first normative standards for a pinch gauge. More recent normative standards have been reported (Mathiowetz et al., 1985; Mathiowetz et al., 1986; Surrey et al., 2001). Researchers have not documented normative data or inter-instrument reliability for all available pinch gauges (Fess, 1986; Su et al., 1995). Researchers have conducted more studies on grip strength instruments than on pinch strength instruments (Fess, 1986; Su et al., 1995).

Kellor (1971) published normative data for the Osco pinch meter. This pinch gauge is no longer available commercially (Mathiowetz et al., 1984). Mathiowetz (1991) found that the Osco pinch gauge was not interchangeable with the B&L pinch gauge. The B&L pinch gauge resulted in higher measurements than the Osco pinch gauge (Mathiowetz et al., 1985). This could be due to Mathiowetz (1991) using three trials per person and Kellor (1971) using two trials per person. Instruments need to have their own normative data when interchangeability does not occur (Mathiowetz 2000). Due to the B&L and Osco instruments not being interchangeable, researchers have established normative data for the B&L pinch gauge. B&L pinch gauge normative data has been established for children ages 5 to 12 (Ager, Olivett, & Johnson 1984;

Surrey et al., 2001), age 6 to 19 (Mathiowetz et al., 1986), and for adults age 20 to 75+ (Mathiowetz et al., 1985).

Researchers establishing normative data for pinch strength in children ages 5 to 12 used the Preston pinch gauge in Ager, Olivett and Johnson (1984) and the B&L pinch gauge in Surrey et al., (2001). Comparing the results gathered from Ager, Olivett, and Johnson (1984) to Surrey et al. (2001) we found that Ager, Olivett, and Johnson (1984) had higher pinch strength scores in most age groups. Both studies had limitations which involved gathering participants from only one geographic location, and unbalanced demographics (i.e. male to female ratio, number of participants in each age group). Surrey et al. (2001) tested participants in highly stimulating environment while Ager, Olivett, and Johnson (1984) tested in a quiet, non-distracting environment. Another possibility for differences in results could be from Ager, Olivett, and Johnson (1984) using only one trial per person for the three pinch patterns whereas Surrey et al. (2001) used three trials per person. Both studies assessed lateral and three-jaw chuck, but Surrey et al. (2001) used pad-to-pad pinch while Ager, Olivett, and Johnson (1984) used tip pinch. These differences could have contributed to the differences in outcomes.

Mathiowetz, Vizenor, and Melander (2000) compared the Baseline hydraulic (0-60 lbs) and Baseline hydraulic (0-50 lbs) pinch gauges against the B&L pinch gauge for inter-instrument reliability and concurrent validity. Participants, age 20-55, included 40 males and 40 females. Only one hand was measured for each instrument tested to reduce possible fatigue. The Baseline hydraulic (0-50 lbs) pinch gauge was tested by only males and the Baseline hydraulic (0-60 lbs) pinch gauge was tested by females. Both genders were tested with the B&L pinch gauge. The Baseline hydraulic (0-60 lbs) and Baseline hydraulic (0-50 lbs) pinch gauges were found to not be interchangeable with the B&L pinch gauge. The differences in pinch means between the

Baseline hydraulic (0-50 lbs) and B&L averaged 8.7% for key pinch and 11.1% for three jaw chuck (Mathiowetz et al., 2000). Tip pinch had a smaller average difference of 5.5%, but the was less than 0.90 (Mathiowetz et al., 2000). The Baseline hydraulic (0-60 lbs) pinch gauge had much larger differences when compared to the B&L (Mathiowetz et al., 2000). Differences averaged 36.6% for tip pinch, 25.8% for key pinch, and 26.8% for three jaw chuck pinch (Mathiowetz et al., 2000). The study did not report which gauge was more accurate (Mathiowetz et al., 2000). The differences found between the pinch gauges could be due to the infrequent calibration checks and whether checks were accurate (Mathiowetz et al., 2000). Mathiowetz et al. (1984) tested the accuracy of the B&L and Preston pinch gauge by hanging known weights from the finger groove of each. When the B&L pinch gauge was compared to the Preston pinch gauge more accuracy was found in the B&L (±1%) (Mathiowetz et al., 1984).

Pinch strength has a correlation with age across the lifespan. Between ages 5 to 20 pinch strength increases with age (Surrey et al., 2001; Ager et al., 1984; Mathiowetz et al., 1986). Scores remained similar from age 20 to 59, with a decline in pinch strength from ages 60-79 (Mathiowetz et al., 1984). Researchers have found hand strength peaks between 25 and 50 years old (Mathiowetz et al., 1984). Ranganathan, Siemionow, Sahgal, and Yue (2001) compared younger subjects (age 20-35) to older subjects (age 65-79). The study found that there was a significant reduction in pinch strength in the older subjects (Ranganathan et al., 2001). When comparing genders, women had more of decline in pinch strength compared to men (Ranganathan et al., 2001). Su et al., 1995 documented pinch strength in 356 adults ages 20 to 88. They found that pinch strength remained stable in men up to age 59 before declines began. After the age of 70 deterioration of pinch strength is more significant (Su et al., 1995). Puh (2010) tested pinch strength in 199 adults age 20 to 79. Puh (2010) found that peak pinch strength occurred between ages 35 to 49 in men and between ages 20-49 in women. Ages 65 to 79 had the weakest pinch strength measurements among age groups (Puh, 2010). These findings are consistent with Shiffman (1992) who found that declines in functional ability of the hand started at age 65 and by 75 the deterioration was more apparent. In all studies mentioned, males were stronger than females across the lifespan. Looking at individuals through their lifespan, their pinch strength decline does not occur as quickly when compared to their decline in grip strength (Su et al., 1995).

There are nine pinch gauges currently on the market. They are either hydraulic or digital. Some of these pinch gauges offer various output readings (every 0.25lb, 0.5lb, 1lb, or 2lb). Pinch gauges also have differences in the maximal amount of force they read, in pounds or kilograms, when assessing pinch strength. It may be easier to record accurate readings with the digital gauges than with the hydraulic gauges because the digital pinch gauges have LCD screens that display the pinch force. However, one cannot assume that just because these instruments are digital they are reliable (Fess, 1986).

No comparison research has been done on the Baseline digital pinch gauge. The purpose of this study is to determine if there is a difference in readings between the Baseline digital pinch gauge and the B&L hydraulic pinch gauge. The B&L pinch gauge was chosen due to it being considered the "gold standard" for its high reliability and validity (Mathiowetz et al., 2000). The research question for this study is, what is the degree of inter-instrument reliability and concurrent validity between the Baseline digital pinch gauge and the B&L hydraulic pinch gauge. Normative data that has been published for the B&L pinch gauge can be generalized to the Baseline digital pinch gauge if no differences are found between the two instruments.

Methods

Participants

Participants were self-reported to be healthy individuals who had no neuromuscular or orthopedic conditions that affected their upper extremity function. Self reported age, gender, hand dominance, and ethnicity were recorded prior to data collection. Recruitment included 54 males and 48 female participants between the ages of 18 and 50. All participants were English speaking, and had no contraindications to performing the pinch strength assessment. Recruitment took place in Toledo, Ohio. Recruitment sites for participants were at the University of Toledo. Cognition was not tested, but individuals were observed to be able to understand instructions and give consent.

Instruments

Instruments used in this study were the Baseline digital pinch gauge and the B&L hydraulic (0- 60 lb) pinch gauge (Figure 1). This study is part of a larger study, which also compares the inter-instrument reliability and validity between the Jamar hydraulic dynamometer and the Baseline digital dynamometer.

Procedures

The University of Toledo Institutional Review board approved the protocols and informed consent document for this study. All participants underwent informed consent prior to participating in our study. Researchers randomized the presentation of the two pinch gauges and two dynamometers to control for order effect. We tested three pinch patterns: lateral, three-jaw chuck, and tip pinch. Lateral pinch involves pressing the pinch gauge between the pad of the thumb and lateral aspect of the index finger (Ager et al. 1984). Three-jaw chuck requires pressing the pinch gauge with the pad of the thumb and the pads of the index and middle fingers (Surrey et al. 2001). Tip pinch involves pressing the pinch gauge between the tip of the thumb and tip of the index finger (Ager et al, 1984). The sequence of the three pinch patterns, and grip strength measurements was randomized between participants to control for order effect. A screen shot of the program used to collect data can be seen in Figure 2.

During testing, participants sat at a designated table and chair. The table and chair were at standard heights and the chair did not have arm rests. The researcher explained the purpose of the study and risks that may be involved to the participant. The researcher determined if the participant met the inclusion criteria and recorded the demographic information.

Researchers used the American Society of Hand Therapist (1992) protocol for evaluating pinch strength in this study. Researchers took three successive trials of each pinch pattern on the left hand during the assessment. This gave a total of 18 trials across the three types of pinch; nine using the B&L pinch gauge and nine using the Baseline digital pinch gauge. There was a 15 second rest period between each trial and a one minute rest period between instruments to control for fatigue. Participant pinch strength testing occurred in a comfortable seated position, with his/her shoulder adducted. The participant’s elbow was flexed at 90 degrees with the forearm and wrist in the neutral position. Wrist extension of 0-30 degrees is allowed during maximal pinch strength measurement (American Society, 1992). The researcher held the distal end of the pinch gauges while the test was administered (Fabrication Enterprises, 2008). We used verbal instructions given to the client on how to perform the pinch patterns per the Mathiowetz, Weber, Volland, and Kashman (1984) study. These instructions include the researcher explaining and demonstrating the proper pinch pattern followed by instructing the participants to then pinch as hard as they can when they are ready. As the participant pinches, the research then says, “Harder!... Harder!... Relax.” Researchers repeated these instructions for the second and third trials (Mathiowetz et al., 1984).

No research reports on how often instrument calibration should be done (Mathiowetz, Vizenor, & Melander, 2000). For our study, the vendor completed calibration of instruments prior to, halfway, and post-study. This was done to increase intra-instrument reliability. The vendor certified that the calibration/performance testing of the instruments was accomplished in accordance with applicable specifications using test weights or equipment periodically calibrated with weights traceable to the National Institutes of Standards and Technology. All instruments were "appropriately calibrated", although any differences in instrument readouts were not relayed to us from the vendor.

Study Design and Data Analysis

This was a cross sectional descriptive study using a randomized crossover design with repeated measures. This type of design protects against type one and type two error if there is no sequence or order effect. Descriptive statistics used in this study are the mean values, range of scores, percent differences between instrument’s mean values, standard deviations (SD) of pinch strength measurements, and the two-way, mixed model intraclass correlation coefficient (ICC). Percent differences were calculated to determine practical significance of any differences between the means (i.e., difference between means divided by the average mean) (Mathiowetz et al., 2000). Effect size was calculated to estimate the magnitude of the relationship regarding the mean differences between instruments (Portney & Watkins, 2009). The Bland-Altman analyses were used to graphically show the average of the differences between the means for each instrument. T–tests for paired measures and the ICC were used to test for inter-instrument reliability and concurrent validity. The level of significance for all analyses was set at an alpha value of less than 0.05.

Results

Demographics

Among the 102 participants involved in the study, ages ranged from 18-50 years. The mean age was 25.51 years, standard deviation was 7.25 years, and median was 23 years. When grouping by age, a majority of the participants were between 21-24 years old (58%). When distributed among gender 53% were male. Hand dominance presented with 88 participants (86.3%) being right-handed, 11 participants (10.8%) left-handed, and 3 participants (2.9%) reported not having a hand dominance. Ethnic backgrounds of participants were as follows: Caucasian (92%), African American (6%), Hispanic (3%), and Asian (1%).

Data Analysis

The means of the three trials for each pinch pattern were used as the outcome measures. Calculations of the mean, mean percent difference, standard deviation, minimum, maximum, and range can be found in Table 1.

Analyses indicated that there were no significant order effects for the paired data. Data analysis using a two-way mixed effects model ICC demonstrated significant correlations between the two pinch gauges of 0.95 for three-jaw chuck, 0.97 for lateral pinch, and 0.86 for tip pinch. The effect size for the three-jaw chuck was 0.28, lateral pinch was 0.19, and tip pinch was 0.54. Comparison of means between instruments using t-tests for paired measures (see Table 2) found statistically significant differences between the B&L and Baseline digital pinch gauge for all pinch patterns ($p < 0.05$). Percent differences between the two instruments were: 7.12% for three-jaw chuck, 4.63% for lateral pinch, and 14.72% for tip pinch.

The Bland-Altman analyses were used to visually show the results of our study. The Bland-Altman Plot is a way to graph the comparison of the two instruments' mean differences for each pinch. This plot interprets the data clinically, not statistically. The further from zero each represented participant's mean average difference is, the greater the mean difference between instruments. The bias line is the average difference between all the mean differences. Figure 3 shows that the three-jaw chuck pinch mean differences between the instruments had a bias of 1.18 pounds. The Baseline digital pinch gauge measured lower pinch forces for the three-jaw chuck compared to the B&L 76 times out of the 102 participants. Figure 4 shows the Bland-Altman Plot for the lateral pinch. A bias of 0.89 pounds was found for lateral pinch. The Baseline digital pinch gauge measured 75 of the 102 participants lower than the B&L for the lateral pinch. Figure 5 shows the Bland-Altman Plot for the tip pinch. Tip pinch demonstrated a bias of 1.76 pounds between instruments. For the tip pinch, the Baseline measured lower for 82 out of the 102 participants compared to the B&L pinch gauge.

Discussion

Often times, clinicians use new technologies and instruments with the assumption that these items have been tested for reliability and validity. However, the Baseline digital pinch gauge has not been tested for its psychometric properties. Therefore, the purpose of this study was to determine if there was a difference in readings between the Baseline digital pinch gauge and the B&L hydraulic pinch gauge. Several of the statistical analyses showed that there were significant differences between the two types of pinch gauges.

Significant findings of the t-test demonstrate that concurrent validity did not occur between the B&L hydraulic and Baseline digital pinch gauge (all p-values < 0.05). The effect size was calculated to complement the p-values obtained and describe the magnitude of differences between all three pinch types. The effect size for the three pinch types ranged from 0.19 to 0.54. The larger the effect size, the greater the difference is between the instruments. A small effect is defined as a d-value of 0.20, moderate effect as 0.5, and a large effect as 0.8 (Cohen, 1988). Both lateral and three-jaw chuck pinch demonstrated small effect sizes, signifying smaller statistical differences between the two instruments. The finding of a moderate effect size with tip pinch indicates that the differences between the gauges are larger for the tip pinch. The larger effect size for the tip pinch may also be due to the mean tip pinch measurements being smaller in comparison to forces seen with the other two pinch patterns. With tip pinch, the mean measurement of 11.01 (Baseline) and 12.67 pounds (B&L) compared to the lateral and three-jaw chuck pinch types at 15.99 (Baseline) to 19.67 pounds (B&L). When a patient is only pinching 11 pounds of force a difference of one to two pounds between instruments is quite substantial, especially when documenting progress in therapy.

A mean percentage difference greater than 5.6% was considered to be clinically significant based on a previous study by Mathiowetz et al. (1985). Percent differences were smallest with the lateral pinch at 4.63% (Table 2), which is not considered clinically significant. Three-jaw chuck had a 7.12% mean difference and tip pinch had a 14.72% mean difference between instruments. These differences are clinically meaningful and add further evidence that the Baseline digital pinch gauge is not interchangeable with the B&L hydraulic pinch gauge.

The results of this study were similar to the findings Mathiowetz et al. reported (2000). Mathiowetz et al. compared the Baseline Hydraulic (0-50lb) pinch gauge to the B&L Hydraulic pinch gauge and found 5.5% to 11.1% mean differences among pinch patterns. Mathiowetz et al. (2000) also compared the Baseline Hydraulic (0-60lb) pinch gauge to the B&L hydraulic pinch gauge and found even greater mean differences (25.8% to 36.6%) between the instruments. Although the instruments in our study demonstrated inter-instrument reliability, mean differences of 7.12% and 14.81% between two of the three pinch patterns are clinically significant.

Upon visual analysis of the Bland Altman plots for the 3-jaw chuck and tip pinch, even variability is seen across averages and mean differences throughout the range of measurements. Figure 4 shows the Bland-Altman Plot for the lateral pinch. More outliers were seen for lateral pinch compared to the other pinches, which remained evenly distributed across the range. We are unsure of the cause, but there was a smaller bias (0.89) for the lateral pinch in comparison to the other two pinch patterns. The Bland Altman Plot demonstrated the Baseline digital pinch gauge is measuring 1.18 pounds lower than the B&L pinch gauge, on average, for three-jaw chuck pinch (Figure 3). In terms of clinical practicality, 1.18 pounds demonstrates a meaningful difference between instruments. Tip pinch showed the highest bias of 1.76 pounds. No trends were seen among the three pinch patterns for mean pinch force differences across the range of participants as shown on the Bland Altman Plots. However, the Baseline digital pinch gauge measured lower pinch forces in comparison to the B&L pinch gauge for a majority of the participants. An average of 78 out of the 102 participants produced lower pinch forces with the Baseline digital pinch gauge across the three pinch patterns tested in comparison to the B&L pinch gauge.

Portney and Watkins, (2009) state that ICC values above 0.75 indicate good reliability, but that for clinical measurements correlations should exceed 0.90. The ICC values in our study ranged from 0.86 to 0.97, which demonstrate good reliability between instruments, but the tip pinch did not achieve a 0.90 ICC needed for clinical measurements. The ICC results for tip pinch were similar to the Mathiowetz et al. (1984) and (2000) studies, demonstrating that it was the least reliable of the three pinches. A strength of our study is shown with the similar variability of standard deviations of the mean between instruments and like pinches, demonstrating consistency in our methods (see Table 1). Proper calibration and use of standardized positioning and instructions were followed per Mathiowetz et al. (1984) recommendations to further reduce variability in this study.

In this study inter-instrument reliability was found, but concurrent validity was not supported. Therefore, the results of this study indicate that the Baseline digital pinch gauge cannot be interchanged with the B&L hydraulic pinch gauge. Failure to demonstrate concurrent validity may be due to a couple factors. Differences in forces produced between the two instruments may be due to the differences in design. For example, the Baseline digital pinch gauge has a concave circular finger surface, whereas the B&L has a more convex rectangular finger surface. Finger size may affect ability to pinch properly on the instrument. We observed that men, in particular, had a difficult time getting their fingers placed correctly on the Baseline digital pinch gauge. Specifically, men with larger fingers had difficulty placing their fingers properly for the 3-jaw chuck pinches. Although the pinch instruments demonstrated inter-instrument reliability, the B&L normative standards should not be used when assessing pinch with the Baseline digital pinch gauge due to not demonstrating concurrent validity.

Implications for Occupational Therapy

Measuring pinch strength is a common assessment used in occupational therapy (Puh, 2010). Occupational therapists measure pinch strength to show improvements or deteriorations following interventions and measurements can be compared to normative standards of individuals of the same sex and age (Brorson, Werner, & Thorngren, 1989; Federman, 1986; Kellor et al., 1971; Mathiowetz, Vizenor, & Melander, 2000; Puh, 2010). Reliable and valid measurements are needed in order to accurately monitor progress (Fess, 1986). Occupational therapists should use of proper body positioning and consistent verbal instructions in clinical practice to enhance reliability of measurements (Mathiowetz et al., 1984; Mathiowetz et al., 1986). Having the client sitting upright, with elbows flexed at 90 degrees, and wrists and forearms in neutral have been shown to be the most reliable ways to measure pinch strength and should be used when evaluating clients. We recommend that occupational therapists use the average of three trials instead of one trial or choosing the highest score, as research demonstrates it to be most reliable and accurate (Mathiowetz et al., 1984). Clinicians should also regularly calibrate instruments to ensure accuracy in measurements, yet no recommendations on a specific recalibration timeframe have been reported (Mathiowetz et al., 1985). Therefore, we recommended that occupational therapists use their best clinical judgment as to timeframe, but have their pinch gauges recalibrated regularly.

We used the Bland Altman Analyses to examine the clinical significance for the mean differences seen between the Baseline digital pinch gauge and the B&L hydraulic pinch gauge. For clinical practice, an average pinch difference of 1.18 to 1.76 pounds of force is a meaningful difference, especially among populations with upper extremity conditions. Percent differences discussed earlier provide further evidence regarding the clinically significant differences between the forces measured by the B&L and Baseline pinch gauges. Both the tip and three-jaw chuck pinch pattern differences were found to be clinically significant, which is similar to a prior study by Mathiowetz et al. (2000). Along with the data demonstrating that concurrent validity did not occur in any pinches, we recommend that occupational therapists not interchange these pinch instruments, especially on the same client, for occupational therapy evaluations. Furthermore, we do not advise that occupational therapists use the normative standards created for the B&L hydraulic pinch gauge with the Baseline digital pinch gauge.

This study sampled individuals without orthopedic conditions and therefore they produced higher pinch forces than may be seen in clinical settings with persons who have hand or upper extremity conditions. Assuming these results can be generalized to adults with orthopedic conditions, we presume that effect sizes may be even larger between instruments due to decreased pinch strength among those with hand conditions. For example, when seeing a patient with carpal tunnel syndrome who is averaging a lateral pinch force of five pounds, a two-pound mean difference between pinch gauge instrument readings would be clinically significant. Given the results of this study, occupational therapists need to be aware that in most cases, lower pinch forces were measured by the Baseline digital pinch gauge than the B&L pinch gauge.

Limitations

A limitation to the study was not being able to precisely measure the intra-instrument reliability. It is unknown whether the instruments remained calibrated for all subjects, however, the calibration was verified by the vendor prior to and mid-way through data collection. Other limitations in this study were that a majority of the participants were between the ages of 21-24 and Caucasian, leading to a lack of age and ethnic diversity among participants. This study also did not evaluate gender differences with pinch strength between the two instruments; however, the sample was relatively evenly divided between men and women participants.

Future Research

Future research is needed to follow up with a more even age distribution and wider variety of ethnicity. Future research is also needed to create normative standards by gender for the Baseline digital pinch gauge.

Acknowledgements

I greatly appreciate my research advisor, Julie Jepsen Thomas, Ph.D., OTR/L, FAOTA, for her support, guidance, assistance, encouragement, and time. I would also like to thank Martin Rice, Ph.D., OTR/L, FAOTA, for his technical skills for creating the randomization program I used for data collection and continual support for this study. Last, but not least, I would like to thank my research partner, Heather Deaton, BS, OTD student, for her support, collaboration, and communication during this study.

References

Ager, C. L., Olivett, B. L., & Johnson, C. L. (1984). Grasp and pinch strength in children 5 to 12 years old. *The American Journal of Occupational Therapy, 38*(2), 107-113.

American Society of Hand Therapists. (1992). *Clinical assessment recommendations*. 2nd ed. Chicago, IL: The Society.

Brorson, H., Werner, C. & Thorngren, K. (1989). Normal pinch strength. *Acta Orthopaedica Scandinavica, 60*(1), 66-68.

Fabrication Enterprises Incorporated (2008). Hydraulic pinch gauges. *Baseline Evaluation Incorporated.* Retrieved from:http://www.fabricationenterprises.com/mm5/graphics/2008CAT/2008CATPDFS/FAB%20 2008%20CATALOG%20000-186-LR%2059.pdf

Fess, E. E. (1986). The need for reliability and validity in hand assessment instruments. *The Journal of Hand Surgery, 11A*(5), 621-623.

Flood-Joy, M. M., & Mathiowetz, V. (1987). Grip strength measurement: A comparison of three Jamar dynamometers. *Occupational Therapy Journal of Research, 7*(4), 235-243.

Hinson, M., Woodard, J., & Gench, B. (1990). Reliability of the Jamar digital dynamometer model 2A. *Occupational Therapy Journal of Research, 10*(2), 108-110.

Kellor, M., Frost, J. H., Siberberg, N., Iverson, I., & Cummings, R. (1971). Hand strength and dexterity. *The American Journal of Occupational Therapy, 25*, 77-83.

Kielhofner, G. (2006). Developing and evaluating quantitative data collection instruments. In G.

Kielhofner (Ed.), *Research in occupational therapy* (pp.155-176). Philadelphia: F. A. Davis Company.

Mathiowetz, V. (1991). Reliability and validity of grip and pinch strength measurements. *Critical Reviews in Physical and Rehabilitation Medicine, 2*, 201-212.

Mathiowetz, V., Kashman, N., Volland, G., Weber, K., Dowe, M., & Rogers, S. (1985). Grip and pinch strength: Normative data for adults. *Archives of Physical Medicine and Rehabilitation, 66*(2), 69-75.

Mathiowetz, V., Vizenor, L., & Melander, D. (2000). Comparison of Baseline instruments to the Jamar dynamometer and the B&L engineering pinch gauge. *The Occupational Therapy Journal of Research, 20*(3), 147-162.

Mathiowetz, V., Weber, K., Volland, G., & Kashman, N. (1984). Reliability and validity of grip and pinch strength evaluations. *The Journal of Hand Surgery, 9*(2), 222-226.

Mathiowetz, V. Wiemer D. M., & Federman, S. M. (1986). Grip and pinch strength: Norms for 6- to 19-year-olds. *The American Journal of Occupational Therapy, 40*(10),705-711.

Niebuhr, B. R., Marion, R., & Fike, L. M. (1994). Reliability of grip strength assessment with the computerized Jamar dynamometer. *Occupational Therapy Journal of Research, 14*(1), 3-18.

Portney, L. F., & Watkins, M. P. (2009). *Foundations of clinical research: Applications to practice, 3ed.* Pearson Education, Inc., New Jersey.

Puh, U. (2010). Age-related & sex-related differences in hand and pinch grip strength in adults. *International Journal of Rehabilitation Research; 33*(1), 4-11. doi: 10.1097/MRR.0b013e328325a8ba

Ranganathan, V. K., Siemionow, V., Sahgal, V., & Yue, G. H. (2001). Effects of aging on hand function. *Journal of the American Geriatrics Society, 49*(11).

Rice, M. S., Leonard, C., & Carter, M. (1998). Grip strengths and required forces in accessing everyday containers in a normal population. *The American Journal of Occupational Therapy, 52*(8), 621-626.

Shiffman, L. M. (1992). Effects of again on adult hand function. *The American Journal of Occupational Therapy, 46*(9), 785-792. doi:10.5014/ajot.46.9.785

Smith, R.O., & Benge, M. W. (1985). Pinch and grasp strength: Standardization of terminology and protocol. *The American Journal of Occupational Therapy, 39*(8), 531-535.

Su, C., Chein, T., Cheng., K., & Su, C. (1995). A study of pinch strength in normal Taiwanese adults. *The Kaohsiung Journal of Medical Sciences, 11*(2); 69-78.

Surrey, L. R., Hodson, J., Robinson, E., Schmidt, S., Schulhof, J., Stoll, L., & Wilson-Diekhoff, N. (2001). Pinch strength norms for 5- to 12- year-olds. *Physical & Occupational Therapy in Pediatrics, 21*(1), 37-49.

Table 1

Descriptive Statistics for Pinch Instruments

		Mean	Percent Difference	Standard Deviation	Min	Max	Range
3 Jaw	Baseline	15.99	7.12	4.69	9.00	30.33	21.33
	B&L	17.17		4.20	10.67	31.33	20.67
Lateral	Baseline	18.78	4.63	5.13	11.00	33.67	22.67
	B&L	19.67		4.67	12.33	31.00	18.67
Tip	Baseline	11.01	14.72	3.26	4.67	19.67	15.00
	B&L	12.76		3.28	6.67	22.67	16.00

Note: N=102

All statistics are in pounds

Table 2

T-Test - Paired Samples Statistics

		Mean	Standard Deviation	Standard Error of the Mean	t	*p value* (2-tailed)
3Jaw	Baseline	15.99	4.69	0.46	-6.97	0.000*
	B&L	17.17	4.20	0.42		
Lateral	Baseline	18.78	5.13	0.51	-5.67	0.000*
	B&L	19.67	4.67	0.46		
Tip	Baseline	11.01	3.26	0.32	-10.26	0.000*
	B&L	12.76	3.28	0.33		

Note: N= 102

*p < 0.05

Figure 1. Instruments used in this study. B&L pinch gauge shown on the left. Baseline digital hydraulic pinch gauge shown on the right.

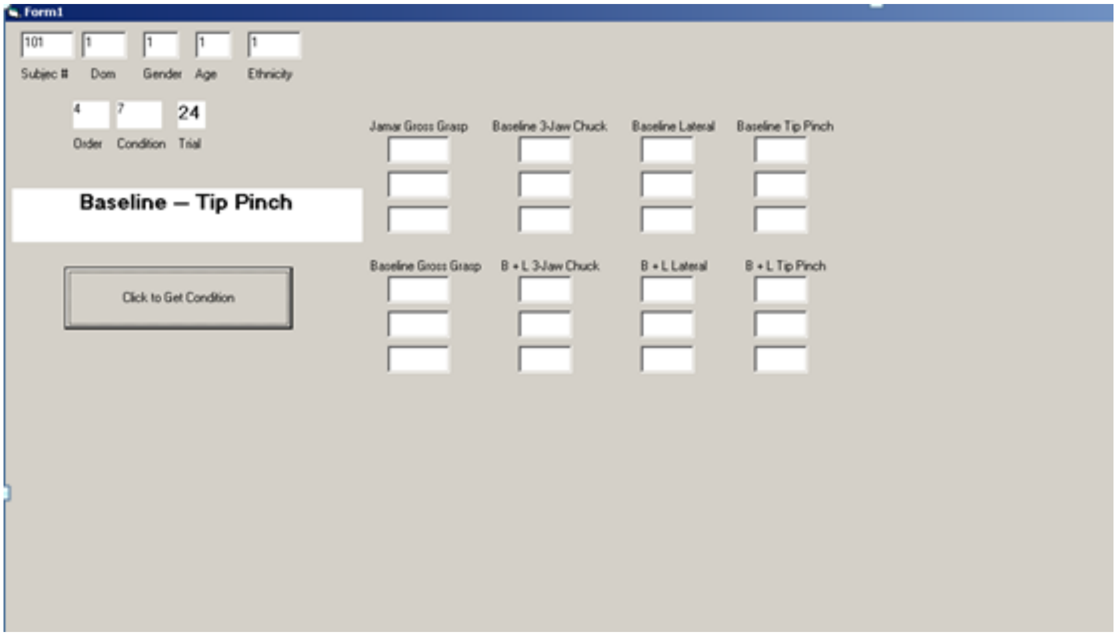

Figure 2. Screen shot of program used to randomize order of conditions and record data.

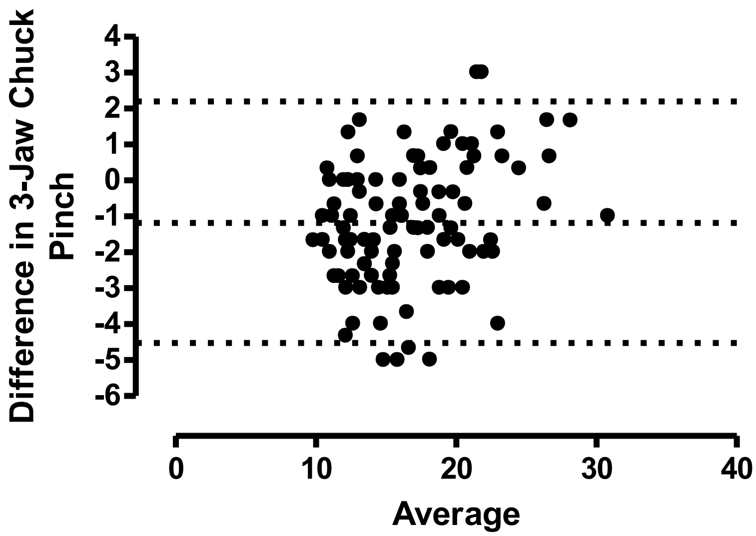

Figure 3. Bland-Altman Plot for 3-Jaw Chuck pinch. The difference of mean averages of each participant represents one dot (102 total). Bias=-1.18; SD=1.70. 95%; confidence intervals (CI) shown in top and bottom dashed lines (-4.52, 2.17). Most participants fell between the confidence interval and no trends were seen among differences in mean averages.

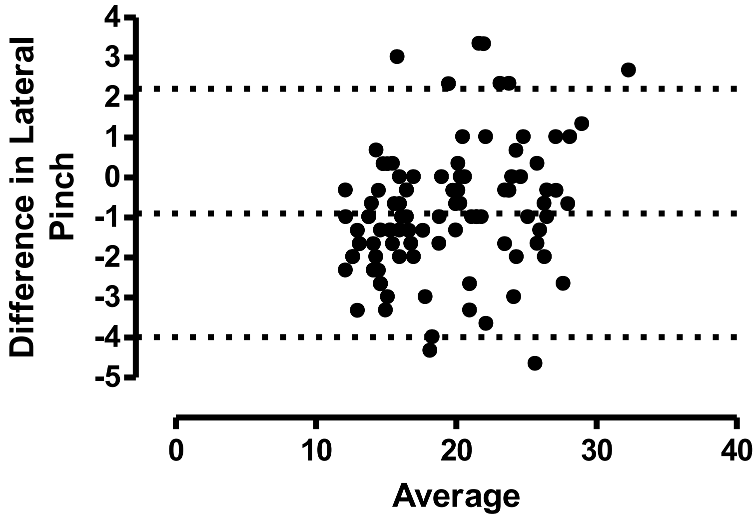

Figure 4. Bland-Altman Plot for Lateral Pinch. Bias = -0.89; SD = 1.58; CI = -3.99, 2.22). No trends were seen among averages, but more data points were outside of the 95% CI.

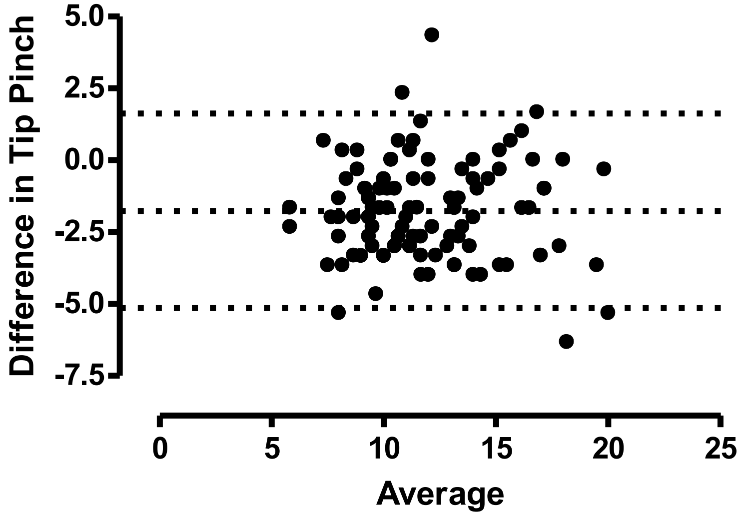

Figure 5. Bland-Altman Plot for Tip Pinch. Bias = -1.76, SD = 1.73, CI= 5.15, 1.63). No trends seen when looking across averages.