EXPERIMENT 3

CHAPTER 8

RELIABILITY OF STRENGTH MEASURES FOR THE DOMINANT AND NON-DOMINANT HANDS

The reliability of a measure refers to its degree of consistency or stability. Test-retest reliability involves administering the same test twice (Voyer, 1998). A perfectly reliable test would give the same results on both occasions. Intraclass correlation coefficients indicate the degree of reliability for continuous measures (Portney & Watkins, 2000). The reliability of grip strength measurements of hands needs to be ascertained so that therapists can determine if a change in grip strength is a real change due to intervention, disease resolution, or progression, or simply within the range of normal measurement error and physiological fluctuation. The reliability of measures of healthy hands needs to be determined before injured hands can be examined. This is because there are many more variables that impact upon injured hands such as fluctuating pain, medication levels, pathology and the psychological and physical responses of the individual to the injury, or disease.

Research has demonstrated that the maximal voluntary isometric grip strength (grip strength) of the uninjured hand is not perfectly reliable over time (Dunwoody, Tittmar & McClean, 1996; Niebuhr, Marion & Fike, 1994; Reddon, Stefanyk, Gill & Renney, 1985) and is influenced by such factors as bed rest, (Greenleaf, Van Beaumont, Convertino & Starr, 1983) time of day (McGarvey, Morrey, Askew & An, 1984; Pearson et al., 1982), medications (Wright, 1959), the expectations of and the gender of the examiner (Rikli, 1974), posture (Fong & Ng, 2001; Watanabe, Owashi, Kanauchi, Mura, Takahara & Ogino, 2005) and other less obvious factors, such as the presence or absence of a depressive psychological state (Emerson, Harrison, Everhart & Williamson, 2001). For women, it is not influenced by their menstrual cycle phase (De Jonge, Boot, Thom, Ruell & Thompson, 2001).

Comparisons between studies

Many hand strength reliability studies have been conducted, but it is difficult to compare them as they have used a variety of isometric or 'dynamic' instruments, age groups, pathologies and retest time intervals, with one or both genders. The type of statistics used to analyse the data has also varied, some groups using Pearson's *r*, others using *ICC*s. In these types of reliability studies when *r* values are given they can be considered to be more liberal estimates of reliability than *ICC* values, refer to Chapter 5 for further details. Thus comparisons between the reliability studies that used either *r* or *ICC* values need to be made with this in mind.

The level of reliability of grip strength dynamometers also needs to be considered. Others have confirmed the excellent reliability of the measurements generated by digital-display dynamometers, or computer-linked Jamar^{TM}-like dynamometers in relation to the analogue Jamar^{TM} (Bohannon, 2005; MacDermid et al., 2001; Niebuhr et al., 1994; Shechtman, Gestewitz et al., 2005; Svens & Lee, 2005).

The factors relating to the definition, assessment method and reliability of handedness have been discussed in Chapter 4. The Edinburgh Handedness Inventory has excellent reliability for the current sample of teenagers (*ICC* (3,1) = 0.9), as demonstrated in Chapter 6.

Many studies have achieved high reliability values ranging from .81 (Wolinsky, Miller, Andresen, Malmstrom & Miller, 2005) to .98 (Shechtman et al., 2003), depending upon the retest time interval, the trial measured (peak, or average), the hand measured (right or left, dominant or non-dominant) and the age of the research participants, with children tending to achieve lower reliability values than adults.

Reliability studies with children

With regard to studies with children older than 10 years, Deutsch and Newell (2001) proposed that age-related enhancements in an isometric finger strength test given to 6, 8 and 10-year-olds were primarily due to a more appropriate, or mature and efficient organization of the sensori-motor system. Researchers in the Netherlands found that optimal performance in terms of force variability control of isometric index finger movements closely matched the maturation of the corticospinal tract up to age 10 (Smits-Engelsman, Westenberg & Duysens, 2003). Thus children over the age of 10 can be expected to give isometric grip strength test performances that are highly reliable over time.

Zverev and Gondwe (2001) retested 41 children aged between 6 and 17 years, at a 2 to 4 day retest interval and achieved a reliability coefficient of .84 (the particular reliability coefficient was not stated). Also the number of children in the sample in their mid to late teenage years was not given. Heeboll-Nielsen (1982) retested 48 children between the ages of 7 and 17 after two hours and achieved a range of non-specified coefficients between .91 to .96, with lower and upper 95% *CI*s of .89 and .97. When examining 18 to 49 year old adults, Schechtman et al. (2003) achieved an *r* value range of .97 to .98.

Influence of gender

The studies that included men and women often did not split the results by gender, possibly due to their small sample sizes. However, when the retest was spread over a number of days or weeks Reddon et al. (1985) found that women on average, achieved slightly higher reliability values than men (.91 versus .94, the particular reliability indices was not stated) and Hamilton et al. (1994) found no significant differences in grip strength values between the genders.

Influence of age

One study examined within-subject variability over six trials with two minute rest breaks, and found that it was highest for the younger children and improved with age in the range of 4 to 16-year-olds (Häger-Ross & Rosbland, 2002). No other studies were found that examined the retest reliability with age as a factor.

Influence of handedness

For adult and childhood studies, the possible influence of handedness on the reliability of grip strength measurements has usually been excluded. Studies have been conducted with only individuals who were right-handed, (Hinson, Woodard & Gench, 1990; Reddon et al., 1985); or they contained a minority of left-handed participants and then combined the results from the dominant hands of the left-handers with those of the dominant hands of the right-handers. Despite this confusion, the two hands appear to behave differently. Several researchers working with adults found that the reliability values for the left hand were slightly higher than those of the right hand (Boissy, Bourbonnais, Carlotti, Gravel & Arsenault, 1999; Bohannon & Schaubert, 2005; Howard & Griffin, 2002; Mathiowetz et al., 1984; Niebuhr et al., 1994) and the reliability of the non-dominant hand to be slightly higher that that of the dominant hand with 2 out of the 29 participants being left-handed (Lagerström & Nordgren, 1998). In contrast, Hinson et al. (1990) who tested 50 15-year-old boys found the reverse to be true. Hamilton et al. (1994) found no significant difference between the hands.

Using a multitude of data base search engines in July 2005 (Journals at OVID, Medline, CINAHL, PsychInfo etc.) only two studies were found that examined the grip strength difference between the two hands when participants were divided into right, mixed and left-handed groups, but these were not reliability studies (Yim et al., 2003; Zverev & Kamadyaapa, 2001). There were no studies located which examined the effect of the degrees of handedness on the reliability of the grip strength measurements of each hand. For example, there were no studies that asked the question "Is the reliability coefficient for the grip strength value of the right hand of a right-hand dominant person higher than that of the right hand of a mixed-handed person?" More details concerning definitions of degrees of handedness and the method for assessing handedness used in this thesis are given in Chapters 4 and 5.

Thus although there have been many grip strength studies over the years, no studies appear to have repeat-tested more than 100 teenagers over the clinically relevant time span of one to four weeks. Hand therapists measure recovery with a number of tools including handgrip dynamometers (Chapter 2). Chapter 7, Part 3 has reported that the grip strength of one hand can be predicted by the grip strength of the contralateral hand; this assumes that the grip strengths of the two hands are stable. Thus it is important to confirm this assumption if prediction equations are to have validity.

The hypotheses

In consideration of the literature on grip strength testing, its reliability, and the possible underlying factors influencing reliability, four hypotheses have been generated and tested here:

(a) The grip strength retest reliability values for both hands would not be significantly higher over a short retest interval of one week, as compared to a longer retest interval of 4 weeks, for the 13 or 17-year-olds, as the physical parameters of the teenagers would not change significantly within four weeks;

(b) Whether tested at a short or long retest interval, the males would generate higher grip strength reliability values, than the females.

(c) Although the effect of gender would result in the males having significantly higher reliability values for the grip strength values of both their hands, than the females, age would not have a significant effect on reliability values;

(d) Those teenagers who were strongly lateralised in their handedness would have greater reliability values for grip strength performance in their dominant hand, than the hands of those whose dominance was not strongly lateralised.

To examine these hypotheses a test-retest protocol was designed to determine the reliability of grip strength in the dominant and non-dominant hands of male and female teenagers, measured over a 1 or 4 week retest interval. Hypothesis d) was developed because the two hands have been shown to be ‘cortically wired’ differently. The dominant hand has been shown to have superior cortical organization and efficiency, as compared to the non-dominant hand (Volkmann et al., 1998). Also a recent study of unilateral arm strength training has demonstrated that for 39 strongly right-handed females strength training only occurs in the right-to-left direction of transfer (Farthing, Chilibeck & Binsted, 2005). Thus those teenagers who were consistently right-handed, or consistently left-handed were expected to have higher reliability values for their dominant hands than the hands of the mixed-handed teenagers.

The information gathered from these experiments is of use to clinicians for assisting in identifying variables that impact on the reliability of grip strength readings from week to week in the remaining healthy hand of unilaterally injured hand therapy patients. The synchronicity of the relationship of the two healthy hands is examined in the next chapter. The behaviour of the injured hand needs to be studied in the future.

METHODS

Selection Criteria for Repeat Testing

All the teenagers who were initially grip strength tested were invited for a second session either within one week, or after four weeks. At the commencement of the second grip strength testing session they were questioned regarding any recent upper limb injuries, or changes in medications (such as Ritalin, for ADHD) that may have occurred since their initial grip strength test session. If there had been injuries or medication changes they were excluded from the retest.

In addition to simply dividing the teenagers into those who were left and those who were right-handed, they were also divided into three handedness groups. These three handedness groups were based on their Laterality Quotients (LQ), taken from their Edinburgh Handedness Inventory scores. The right-handers were defined as those who had LQs between +100 and +70, the mixed-handers had LQs between -60 and +60 and the left-handers had LQs between -70 and -100. The dominant hand of the mixed-handers was taken to be the right hand if the teenager had a positive LQ (between +10 and +60) and was taken to be the left hand if they had a negative LQ (between -10 and -60). Any teenagers with an LQ of zero were asked to report which hand they considered to be their dominant hand. Grip strength testing was always commenced with the dominant hand.

Although research has shown that grip strength is affected by extremes in time of day (McGarvey et al., 1984) but it does not vary significantly from mid-morning to mid-afternoon (Young, Pin, Kraemer, Gould, Nemergut, & Pellowski, 1989). Thus the teenagers were initially and repeat tested between 9.30 am and 3.00 pm.

STATISTICAL ANALYSIS

There is debate as to whether to use the strongest or the average grip strength score from one session to the next in reliability studies. Researchers have found the average of two or three trials to generate slightly higher reliability figures than the strongest force reading (Hamilton et al., 1994; Mathiowetz et al., 1984). The current study used the strongest reading. This was because by using paired samples *t* tests it was found that for each of the four age and gender groups (13-year-old males, 13-year-old females etc.) there was no significant difference between the strongest and the average of three trails on the optimal handle position. The *r* values for retest with the strongest or the average ranged between .949 and .986 for the right hand and between .951 and .988 for the left hand. So it is acceptable to present either the value of the strongest or the average. Using the strongest squeeze is in line with the concept of a "real potential maximum" as done by others (Petersen et al., 1989).

In this current study, with its one examiner and retest design, the Intraclass Correlation Coefficient Model 3,1 [*ICC* (3,1)] was used for calculating reliability of the grip strength scores separately for the dominant and non-dominant hands from T1 to T2 because it was the most appropriate correlation coefficient for this type of experiment (Portney & Watkins, 2000). The Standard Error of Measurement (*SEM*) and their derived Minimum Differences to be Exceeded (*MDE*) were used to ascertain how much change in grip strength could be considered to be real change beyond measurement error. The upper and lower 95% *CI*s of the *ICC*s were examined to determine if any of the groups achieved significantly higher correlations than any other group. Two-way independent groups ANOVAs were performed to evaluate the interaction of combinations of two out of the three variables of age, gender and retest time interval (time interval), as presented below. The absolute percentage changes in the grip strength of the dominant and non-dominant hands were used as separate dependent variables in the ANOVAs. The strengths of any significant main effects were based on the eta-squared values. Eta squared was computed by dividing the sum of squares for each factor, or the interaction of the factors, by the total sum of squares (Tilley, 1999). Strength in this case referred to the percentage of variability in the scores that could be accounted for by the main effect under question. The means of the absolute percentage change were examined to determine the direction of any significant main effects. The level of significance for these tests was set at $p < .05$ and all the analyses were performed using SPSS Version 11.0 for Windows.

RESULTS

Repeat Grip Strength Tested Teenagers

It was intended that all of the 238 teenagers who were initially grip strength tested would be retested. Due to factors such as the study commitments of the teenagers, injuries between the initial grip strength test and the repeat test, early school leavers in Year 12, absenteeism etc., only 154 of the original group were retested.

The repeat test sessions were conducted in the same room with the same methodology as the initial test, but anthropometric variables were not re-measured (see Chapter 7, Part 1a). The retest time interval for any particular teenager was based on their availability. There were valid results for 84 teenagers who were retested within seven days, who constituted the short retest interval group (SIG) and 70 teenagers who were retested after four weeks, who constituted the long retest interval group (LIG), see **Table 8.1**. The tests were scheduled so that the majority of the teenagers were retested at the same time of day and on the same day of the week, see **Table 8.1**. The strongest grip strength of the dominant and non-dominant hands with their *SD*s, have been presented for each age and gender group at the initial session (T1) and at the retest session (T2) in **Table 8.2**.

The use of paired sample *t* tests demonstrated a non-significant difference between the dominant and non-dominant hands from T1 to T2 for each age and gender group ($p < .001$). The dominant hand was, on average, always the stronger one. Note also the general trend at T2 for the mean grip strength to be 0.35 kg f stronger for the dominant hand and 0.75 kg f stronger for the non-dominant hand.

Table 8.1. Sample Numbers and Retest Time Intervals Measured in Days.

Group	Males (*n*)	Females (*n*)	Mean Retest Interval (*SD*)	Range in Days	% Same Time of Day[a]	% Same Day of the Week	% Same Time & Same Day of Week[a]
13 y.o. SIG	22	19	2.49 (3.01)	1 to 9	68	2.4	0
13 y.o. LIG	16	15	27.48 (0.89)	26 to 28	61	80.6	58.1
17 y.o. SIG	18	25	6.56 (0.66)	5 to 7	88	75	70
17 y.o. LIG	20	19	29.20 (4.76)	27 to 49	79.4	94.9	74.3

Note. SIG = short retest interval group; LIG = long retest interval group.
[a]Equal to or less than one hour from the original test time of day.

Table 8.2. Mean (*SD*) Strongest Dominant (Dom) and Non-dominant (Non-dom) Grip Strengths in Kilograms Force From Test 1 (T1) to Test 2 (T2) for Each Age and Gender Group.

Age and Gender	Dom T1 (*SD*)	Dom T2 (*SD*)	Non-dom T1 (*SD*)	Non-dom T2 (*SD*)
13 y.o. males	30.6 (7.2)	31.1 (7.7)	28.7 (6.5)	29.4 (7.3)
13 y.o. females	26.1 (4.2)	26.9 (4.3)	24.0 (4.0)	25.1 (4.2)
17 y.o. males	49.2 (8.4)	49.1 (8.6)	45.7 (8.5)	46.0 (8.4)
17 y.o. females	29.9 (4.6)	30.1 (4.4)	26.7 (4.4)	27.5 (4.0)

Note. Dom = dominant hand; Non-dom = non-dominant hand; T1 = test 1; T2 = test 2.

As can be seen from **Table 8.1**, sample numbers would have been reduced to less than 20 in some groups if divided by age, gender, and time interval. Thus when the results were divided by retest time interval they were sorted into either age, or gender groups, not simultaneously into age and gender groups.

(a) Reliability of the Grip Strength Measurements for the Dominant and Non-dominant Hands with Two Retest Intervals, Split by Age

Although both were statistically significant and excellent, the *ICC*s of the 17-year-olds were higher than those of the 13-year-olds, for both the dominant and non-dominant hands, with generally narrower and higher *ICC* confidence interval values, see **Table 8.3 and 8.4**. The grip strength reliability values of the 17-year-olds tended to have slightly higher *SEM*s and wider *SEM* 95% *CI*s for their time interval and age matched groups. Thus their *MDE*s were also higher. Their variances were also more than double that of the 13-year-olds. Thus they needed greater changes in grip strength before the changes could be considered to be real changes and not just measurement error. The grip strength reliability values of the dominant hands were higher than those of the non-dominant hand, with one exception, the 17-year-old LIGs. For them the grip strength of the non-dominant hand had a slightly higher reliability coefficient with slightly smaller *SEM*s, narrower *SEM* 95% *CI*s and correspondingly smaller *MDE*s.

Table 8.3 Correlation Statistics Split by Age and Time Interval for the Dominant Hand.

Group	*n*	*ICC*	95% *CI* for *ICC*	*SEM*	95% *CI* for *SEM*	*MDE*	Variance
13 y.o. SIG	41	.918	.852 – .956	1.61	1.32 – 2.06	4.46	31.54
13 y.o. LIG	31	.940	.877 – .970	1.81	1.44 – 2.42	5.02	53.83
17 y.o. SIG	43	.969	.944 – .983	2.17	1.91 – 2.94	6.01	155.73
17 y.o. LIG	39	.934	.880 – .965	2.74	2.26 – 3.56	7.61	114.07

Note. SIG = short retest interval group; LIG = long retest interval group; *ICC* = intraclass correlation coefficient (3,1); *CI* = confidence interval; *SEM* = standard error of measurement (kg f); *MDE* = minimum difference to exceed.

Table 8.4. Correlation Statistics Split by Age and Time Interval for the Non-dominant Hand.

Group	*n*	*ICC*	95% *CI* for *ICC*	*SEM*	95% *CI* for *SEM*	*MDE*	Variance
13 y.o. SIG	41	.902	.824 – .947	1.61	1.33 – 2.07	4.47	26.52
13 y.o. LIG	31	.92	.840 – .960	1.97	1.58 – 2.63	5.46	48.54
17 y.o. SIG	43	.96	.927 – .978	2.46	2.04 – 3.15	6.81	151.03
17 y.o. LIG	39	.965	.936 – .982	1.94	1.60 – 2.47	5.37	107.34

Note. SIG = short retest interval group; LIG = long retest interval group; *ICC* = intraclass correlation coefficient (3,1); *CI* = confidence interval; *SEM* = standard error of measurement (kg f); *MDE* = minimum difference to exceed.

ANOVA Results for Age and Time Interval for the Dominant Hand

A two-way ANOVA was conducted using the two age groups and the two time interval groups as the fixed factors, with the absolute percentage change in the grip strength of the **dominant** hand from initial to retest (DomT1 / DomT2 x 100) as the dependent variable. There were no significant main effects of age, time interval, or an interaction effect of the two factors, see **Table 8.5**.

Table 8.5. ANOVA for Time Interval and Age for the Dominant Hand.

Source	Sum of squares	*df*	Mean Square	*F*	*p*
Time interval	77.638	1	77.638	2.257	.135
Age	21.440	1	21.440	0.623	.431
Age x Time	37.850	1	37.850	1.100	.296
Error	5160.144	150	34.401		
Total	5281.775	153			

ANOVA Results for Age and Time Interval for the Non-Dominant Hand

The same two-way ANOVA was conducted using the grip strength of the **non-dominant** hand from initial to retest (Non-DomT1 / Non-DomT2 x 100) as the dependent variable. Again there were no significant main effects for age, time interval, or for the interaction effect of the two, see **Table 8.6**.

Table 8.6. ANOVA for Time Interval and Age for the Non-Dominant Hand

Source	Sum of squares	*df*	Mean Square	*F*	*p*
Time interval	3.516	1	3.516	0.079	.779
Age	16.114	1	16.114	0.363	.548
Age x Time	111.455	1	111.455	2.51	.115
Error	6661.666	150	44.411		
Total	6789.957	153			

(b) Reliability of the Grip Strength Scores of the Dominant and Non-dominant Hands With Two Retest Intervals, Split by Gender

Although all groups achieved significant and good to excellent *ICC* (3,1) values, the 95% confidence intervals for the *ICC*s were consistently higher and narrower for the males (see **Tables 8.7** and **8.8**). For the dominant hand of the males there was a significant but marginal improvement in the ICC if the hand was tested after one rather than four weeks. They demonstrated an excellent *ICC* of .985, with a small *MDE* of 4.53 kg f. For their non-dominant hands there was no significant difference between testing them after one or four weeks. For the females, there was no significant difference in the *ICC*s between the hands, or between the retest time intervals. The *MDE*s were very similar to those of the males but their mean grip strength values were up to 50% less than that of the males. Consequently their absolute percentage changes were greater at 8.0 to 9.7%, as compared to the percentage changes of 5.8 to 6.5% for the males.

Table 8.7. Correlation Statistics and Absolute Mean % (Mean %) Change in Grip Strength Split by Gender and Time Interval for the Dominant Hand.

Group	*n*	*ICC*	95% *CI* for *ICC*	Mean %	*SEM*	95% *CI* for *SEM*	*MDE*	Variance
Male SIG	40	.985	.971 –.992	5.32	1.63	1.37 – 2.15	4.53	177.86
Male LIG	36	.943	.892 – .970	6.48	2.52	2.05 – 3.29	6.97	111.01
Female SIG	44	.786	.639 – .877	7.95	2.11	1.76 – 2.69	5.86	20.86
Female LIG	34	.751	.557 – .868	9.66	2.45	1.99 – 3.24	6.8	24.17

Note. SIG = short retest interval group; LIG = long retest interval group; *ICC* = intraclass correlation coefficient (3,1); *CI* = confidence interval; Mean % = mean % change; *SEM* = standard error of measurement (kg f); *MDE* = minimum difference to exceed.

Table 8.8. Correlation Statistics and Absolute Mean % (Mean %) Change in Grip Strength Split by Gender and Time Interval for the Non-dominant Hand.

Group	*n*	*ICC*	95% *CI* for *ICC*	Mean %	*SEM*	95% *CI* for *SEM*	*MDE*	Variance
Male SIG	40	.976	.956 – .988	6.08	1.97	1.61 – 2.52	5.45	161.24
Male LIG	36	.959	.921 – .979	5.82	1.95	1.59 – 2.55	5.41	92.76
Female SIG	44	.72	.541 – .837	9.55	2.2	1.82 – 2.80	6.1	17.25
Female LIG	34	.809	.651 – .900	8.42	1.99	1.82 – 2.80	5.5	20.64

Note. SIG = short retest interval group; LIG = long retest interval group; *ICC* = intraclass correlation coefficient (3,1); *CI* = confidence interval; Mean % = mean % change; *SEM* = standard error of measurement (kg f); *MDE* = minimum difference to exceed.

Two sets of two-way ANOVAs were conducted for the dominant and non-dominant hands separately, see **Table 8.9** and **8.10**. Using gender and time interval groups as the fixed factors, with the absolute percentage change in the grip strength of either the **dominant** or **non-dominant** hand from initial to retest (DomT1 / DomT2 x 100) as the dependent variable there were no significant main effects of time interval, or an interaction effect of time interval and gender, there was a significant effect of gender. The strength of this effect, based on eta squared, revealed that gender only accounted for 6.2% and 4.7% of the variability in the scores of the dominant and non-dominant hands, respectively.

Table 8.9. ANOVA for Gender and Time Interval for the Dominant Hand.

Source	Sum of squares	*df*	Mean Square	*F*	*p*
Time interval	78.643	1	78.643	2.411	.123
Gender	321.35	1	321.35	9.851	.002
Time x Gender	2.912	1	2.912	0.089	.766
Error	4893.054	150	32.62		
Total	5281.775	153			

Table 8.10. ANOVA for Gender and Time Interval for the Non-dominant Hand.

Source	Sum of squares	*df*	Mean Square	*F*	*p*
Time interval	4.309	1	4.309	0.1	.752
Gender	317.277	1	317.277	7.4	.007
Time x Gender	17.347	1	17.347	0.405	.526
Error	6431.279	150	42.875		
Total	6789.957	153			

(c) Reliability of the Grip Strength Measurements of the Dominant and the Non-Dominant Hand, Split by Age and Gender

Each of the four age and gender groups had reliability values which were significant ($p < .01$) and ranged from good to excellent for both their hands. Within each group there were not significant differences between the reliability values of the grip strength of the two hands, with the scores of the dominant hands achieving slightly higher reliability values for the 13-year-olds; this trend was reversed for the 17-year-olds, see **Tables 8.11** and **8.12**.

Table 8.11. Correlation Statistics and Absolute Mean % (Mean %) Change in Grip Strength Split by Age and Gender for the Dominant Hand.

Group	*n*	*ICC*	95% *CI* for *ICC*	Mean %	*SEM*	95% *CI* for *SEM*	*MDE*	Variance
13 y.o. males	38	.946	.899 – .972	6.31	1.73	1.42 – 2.33	4.79	55.25
13 y.o. females	34	.761	.573 – .873	9.08	2.08	1.68 – 2.75	5.77	18.13
17 y.o. males	38	.91	.833 – .952	5.43	2.53	2.08 – 3.30	7.01	70.99
17 y.o. females	44	.718	.537 – .836	8.4	2.4	1.99 – 3.06	6.65	20.4

Note. SIG = short retest interval group; LIG = long retest interval group; *ICC* = intraclass correlation coefficient (3,1); *CI* = confidence interval; Mean % = mean % change; *SEM* = standard error of measurement (kg f); *MDE* = minimum difference to exceed.

Table 8.12. Correlation Statistics and Absolute Mean % (Mean %) Change in Grip Strength Split by Age and Gender for the Non-dominant Hand.

Group	*n*	*ICC*	95% *CI* for *ICC*	Mean %	*SEM*	95% *CI* for *SEM*	*MDE*	Variance
13 y.o. males	38	.941	.889 – .969	6.08	1.67	1.37 – 2.26	4.63	47.3
13 y.o. females	34	.731	.526 – .856	9.55	2.14	1.72 – 2.81	5.93	17.02
17 y.o. males	38	.924	.859 – .960	5.82	2.32	1.90 – 3.02	6.43	70.81
17 y.o. females	44	.729	.553 – .842	8.42	2.18	1.81 – 2.77	6.05	17.55

Note. SIG = short retest interval group; LIG = long retest interval group; *ICC* = intraclass correlation coefficient (3,1); *CI* = confidence interval; Mean % = mean % change; *SEM* = standard error of measurement (kg f); *MDE* = minimum difference to exceed.

Between some of the age and gender groups there were significant differences in the reliability scores. When the 13-year-old males were compared with the 13-year-old females, the reliability values for the grip strength scores of both hands of the males were significantly greater than that of the matched hands of the females, as judged by the lack of overlap between their *ICC* confidence interval ranges, see **Tables 8.11** and **8.12.** The reliability values for the measurements of both hands for the 17-year-old males were higher than that for the 17-year-old females. However they were only *significantly* higher for the non-dominant hands, as assessed by the lack of overlap in their confidence interval ranges, see **Tables 8.11** and **8.12**.

For both hands, the mean percentage changes were greater for the age-matched females, as compared to their male peers. The *SEM*s and *MDE*s were larger for the 13-year-old females, as compared to the 13-year-old males. These trends were reversed in the 17-year-olds, but by a smaller magnitude. It must be noted that the restricted range of the grip strength values of the females could affect their *ICC* values, however when their *SEM*, *SEM* 95% *CI*s and *MDE* values were compared they were consistent with the reliability picture given by the *ICC*s. The similar trends in the absolute mean percentage changes in grip strength also confirmed this pattern of both hands; namely that the older males had smaller changes in grip strength readings over time.

Two sets of two-way ANOVAs were conducted using the two age and the two gender groups as the fixed factors, with the absolute percentage change in the grip strength of the **dominant** and then the **non-dominant** hand from initial to retest (DomT1 / DomT2 x 100) as the dependent variables, see **Tables 8.13** and **8.14**. There were no significant main effects of age, or interactions effect of age and gender for either hand, but there was a significant effect of gender for both hands. Eta squared revealed that gender only accounted for 6.0 and 5.2% of the variability in the scores for the dominant and non-dominant hands, respectively.

Table 8.13. ANOVA for Age and Gender for the Dominant Hand.

Source	Sum of squares	*df*	Mean Square	*F*	*p*
Age	23.19	1	23.19	0.703	.403
Gender	313.87	1	313.87	9.509	.002
Age x Gender	0.416	1	0.416	0.013	.911
Error	4951.219	150	33.008		
Total	4951.219	150			

Table 8.14. ANOVA for Age and Gender for the Non-Dominant Hand.

Source	Sum of squares	*df*	Mean Square	*F*	*p*
Age	18.501	1	18.501	0.432	.512
Gender	350.732	1	350.732	8.185	.005
Age x Gender	7.207	1	7.207	0.168	.682
Error	6427.224	150	42.848		
Total	6789.957	153			

(d) Reliability of the Grip Strength Measurements Based on Degrees of Handedness

When the teenagers were divided into right, mixed and left-handed groups, the *ICC* (3,1) values of the dominant and non-dominant hands of the left-handers were found to be less reliable, with their *p* values greater than .05, and their 95% *CI*s on either side of zero, see **Tables 8.15** and **8.16**. The left-handers were found to be a small group, so they were excluded from further analyses. When the results for the right-and mixed-handed teenagers were analysed, the dominant hand again tended to generate higher reliability values than the non-dominant hand, with the exception of the mixed-handed teenagers. However none of these *ICC* differences were significant. The *SEM*s, their 95% *CI* and the *MDE*s were similar for the right and mixed-handed groups.

Table 8.15. Correlation Statistics for the Dominant Hands Sorted by Handedness Classifications.

Group	*n*	*ICC*	95% *CI* for *ICC*	*SEM*	95% *CI* for *SEM*	*MDE*	Variance
(R) Dom[a]	86	.966	.948 - .977	2.06	1.81 – 2.44	5.7	124.53
Mixed[a]	58	.959	.931 – .975	2.21	1.89 – 2.74	6.14	119.5
(L) Dom[b]	10	.472	-.182 – .836	3.23	2.27 – 6.03	8.96	19.79

Note. (R) Dom = right-dominant; Mixed = mixed-handed; (L) Dom = left dominant; *ICC* = intraclass correlation coefficient (3,1); *CI* = confidence interval; *SEM* = standard error of measurement (kg f); *MDE* = minimum difference to exceed.
[a]$p < .001$. [b]$p = .072$.

Table 8.16. Correlation Statistics for the Non-dominant Hands Sorted by Handedness Classifications.

Group	*n*	*ICC*	95% *CI* for *ICC*	*SEM*	95% *CI* for *SEM*	*MDE*	Variance
(R) Dom[a]	86	.963	.943 – .975	2.01	1.77 – 2.40	5.58	109.61
Mixed[a]	58	.964	.940 – .978	2.24	1.72 – 2.49	6.2	113.53
(L) Dom[b]	10	.356	-.313 – .789	2.58	1.68 – 4.40	7.14	10.32

Note. (R) Dom = right-dominant; Mixed = mixed-handed; (L) Dom = left dominant; *ICC* = intraclass correlation coefficient (3,1); *CI* = confidence interval; *SEM* = standard error of measurement (kg f); *MDE* = minimum difference to exceed.
[a]$p < .001$. [b]$p = .141$.

When the grip strength values of the right and mixed-handed teenagers were split by gender, all groups had *ICC* values that were above .795 and therefore all achieved excellent reliability ($p < .01$), see **Tables 8.17** and **8.18**. However, the grip strength retest scores of the females where always significantly less reliable than those of the males; as judged by the lack of overlap in the 95% *CI* limits for the females with the males. This held true for the measurements of both hands. The dominant hand of each mixed-handed group had slightly lower reliability values than the dominant hand of the gender-matched, right-handed groups. For the non-dominant hands this pattern held true for the females. However, the non-dominant hand of the mixed-handed males achieved an almost equal *ICC* value to the non-dominant hand of the right-handed males and more desirable *SEM*s and *MDE*s**.**

An examination of the values of the absolute mean percentage changes in grip strength for the above factors revealed that the females consistently had the highest absolute mean percentage change from T1 to T2 in both their hands, see **Tables 8.17** and **8.18**. The hand with the greatest absolute mean percentage change was the non-dominant hand of the mixed-handed females. This was consistent with this group having the lowest *ICC* value.

Table 8.17. Correlation Statistics and Absolute Mean % (Mean %) Change Split by Handedness Classification and Gender for the Dominant Hand.

Group	*n*	*ICC*	95% *CI* for *ICC*	Mean %	*SEM*	95% *CI* for *SEM*	*MDE*	Variance
(R) Dom M	43	.975	.955 – .987	5.95	2.07	1.70 – 2.62	5.72	170.76
(R) Dom F	43	.818	.687 – .897	8.17	2.02	1.70 – 2.59	5.61	22.51
Mixed M	30	.955	.908 – .978	5.29	2.25	1.80 – 3.04	6.28	112.14
Mixed F	28	.796	.606 – .900	8.62	2.01	1.58 – 2.72	5.57	19.79

Note. (R) Dom = right-handed; Mixed = mixed-handed; M = males; F = females; *ICC* = intraclass correlation coefficient (3,1); *CI* = confidence interval; Mean % = mean % change; *SEM* = standard error of measurement (kg f); *MDE* = minimum difference to exceed.
$p < .001$.

Table 8.18. Correlation Statistics and Absolute Mean % (Mean %) Change Split by Handedness Classification and Gender for the Non-dominant Hand.

Group	*n*	*ICC*	95% *CI* for *ICC*	Mean %	*SEM*	95% *CI* for *SEM*	*MDE*	Variance
(R) Dom M	43	.967	.939 – .982	6.97	2.18	1.82 – 2.80	6.04	143.63
(R) Dom F	43	.823	.696 – .900	7.91	1.79	1.49 – 2.29	4.96	18.08
Mixed M	30	.969	.935 – .985	4.85	1.82	1.47 – 2.48	5.03	106.27
Mixed F	28	.761	.543 – .882	9.93	2.1	1.63 – 2.80	5.81	18.38

Note. (R) Dom = right-handed; Mixed = mixed-handed; M = males; F = females; *ICC* = intraclass correlation coefficient (3,1); *CI* = confidence interval; Mean % = mean % change; *SEM* = standard error of measurement (kg f); *MDE* = minimum difference to exceed.
$p < .001$.

Gender and handedness factors were assessed for their impact on reliability using ANOVAs. The following ANOVAs showed that gender had a significant effect for the measurements for both hands, but that handedness classification (right or mixed-handed) and the interaction of the two did not, see **Tables 8.19** and **8.20**. Eta squared indicated that only 5.6% and 5.3% of the variability in the dominant and non-dominant handgrip strength scores could be accounted for by gender, respectively.

Table 8.19. ANOVA for Handedness Group and Gender for the Dominant Hand.

Source	Sum of squares	*df*	Mean Square	*F*	*p*
Handedness	0.376	1	0.376	0.012	.914
Gender	265.979	1	265.979	8.361	.004
Handedness x Gender	10.817	1	10.817	0.34	.561
Error	4453.593	140	31.811		
Total	4720.49	143			

Table 8.20. ANOVA for Handedness Group and Gender for the Non-dominant Hand.

Source	Sum of squares	*df*	Mean Square	*F*	*p*
Handedness	8027	1	8027	0.002	.964
Gender	313.589	1	313.589	7.789	.006
Handedness x Gender	147.685	1	147.685	3.668	.058
Error	5636.592	140	40.261		
Total	6029.485	143			

DISCUSSION

Neither age nor retest time interval had significant effects on the reliability of the grip strength measurements of either hand, within each gender group. Thus the first hypothesis that the grip strength retest correlations for both hands would not be significantly higher over a short retest interval of one week, as compared to 4 weeks, for either age group was supported. Males had significantly higher reliability indices for the grip strength of both hands than females. Further, the second hypothesis that the males would have higher reliability values, whether tested at a short, or long retest interval than the females, was supported. The grip strength measurements of the 13-year-old males were not significantly more reliable than those of the 17-year-old males. Likewise, the third hypothesis that the age of the males would not have a significant effect on their reliability values was supported.

The measurement of the dominant hand of the right-handed teenagers achieved the highest *ICC* (3,1) value with the narrowest 95% *CI*, for the three handedness categories. Thus the fourth hypothesis that those teenagers who were strongly lateralised in their handedness would have greater reliability values for the grip strength performance of their dominant hands, than those whose dominance was not strongly lateralised, was supported. In the small group of left-handers, the measurements for both hands had poor reliability. For both their hands, the lower limits of their 95% *CI*s for their *ICC* values extended into negative values, with their non-dominant hand measurements achieving even lower reliability values (*ICC* = .356) than that of their dominant hands (*ICC* = .472).

Reliability of Grip Strength Test Results in General

There are several aspects of reliability addressed in this thesis. *Instrument reliability* was excellent, as shown by the calibration results of the GripTrack™ as reported in Chapter 5. *Inter-rater reliability* was not an issue as there was only one tester (the author). *Intra-rater reliability* could only be assessed within the context of the reliability of the teenagers' measurements and could not be separately evaluated from the other variables. However, it was expected that *intra-rater reliability* was excellent as the variables in methodology were controlled for by (a) standardizing all test protocol, as discussed in Chapter 5 and (b) by using a tester (the author) who was experienced in the use of the test equipment prior to the data collecting. The other factor to consider was examiner bias. This was reduced by not having the results of T1 consulted, or visually displayed, for each teenager as they were performing T2, one to four weeks later. The tester was thus effectively blinded to the results because it was not possible to spontaneously recall the T1 test results for each of the 154 repeat tested teenagers at the time of their T2 session. In this way the tester was blinded to the results of T1 before T2 was completed, this in turn would have contributed to decreasing examiner bias and examiner expectations (Rikli, 1974).

Variables such as time of day and test conditions were controlled for as much as possible, as described above. *Within-subject reliability* was the only variable which was difficult to control. One aspect of this was motivation, but it was assumed that motivation was high because of the low retest-refusal rate (only two students declined to be retest) and the eagerness of some of the boys to know and compare their grip strength results with their classmates. Their best T1 and T2 test results were displayed on the GripTrack™ programme-generated bar graph after T2 was completed. Each teenager was asked if they thought that they had put in greater or less effort on T2 than on T1. Although not tallied, the general impression was that the boys were concerned that their sincerity of effort was seemingly being questioned, and they had hoped that their grip strength had improved, so as to better the scores of their classmates. The majority of the females demonstrated no such concern, but thought that they had performed consistently from T1 to T2. Although not reaching significance, the absolute mean percentage change in grip strength from T1 to T2 was always positive and ranged from 4.85% (the non-dominant hand of the mixed-handed males) to 9.93% (the non-dominant hand of the mixed-handed females) depending on which age, gender, hand and handedness variables were used to split the data. These changes may have been due to a motor learning effect.

(a) The Reliability of the Grip Strength of the Dominant and Non-Dominant Hands with Short and Long Retest Intervals, Split by Age

When divided by retest time interval and age, the grip strength of the two hands of all the teenagers, displayed good to excellent reliability. The older teenagers and the short retest time interval gave the most reliable repeat strength values for the dominant hand. The *ICC*s of either hand ranged from .902 to .969, whether they were retested at a short or a long retest interval. Although excellent reliability is generally regarded as having *ICC*s higher than .75 (Portney & Watkins, 2000), such qualitative descriptions do not consider the influence of different magnitudes of the variances.

The fact that the dominant hands of the older teenagers at a short retest interval had the highest *ICC* value, but at the same time did not possess the smallest *SEM*s, with the narrowest *SEM* 95% *CI*, or in turn the smallest *MDE*, highlights the disparate information on reliability that these statistics provide. Thus the clinician that is interested in reliability over time should note that the teenagers that had the smallest magnitude of error between tests were the younger teenagers when tested at a one week retest interval. But these 13-year-olds also had narrower grip strength ranges and therefore narrower variance values, which in turn give smaller *SEM*s and smaller *MDE*s (Stratford and Goldsmith, 1997). The slightly lower *ICC*s of the younger teenagers may simply have been a reflection of their lower variances.

The differences in the *ICC* values between the teenagers when they were arranged by age and retest time interval were not significant, as the upper and lower limits of the

95% *CI* all overlapped. Also the ANOVAs demonstrated no significant effect of age, retest time interval, or the interaction of the two. There does not appear to be any other reliability studies with teenagers with which to compare the effects of age and two different retest time intervals.

(b) The Reliability of the Grip Strength of the Dominant and Non-Dominant Hands with Short and Long Retest Intervals, Split by Gender

Comparisons of the *ICC* (3,1) values between the males and the females were hazardous as the total mean squares variance of the measurements from which they derived should always be similar in order to enable fair comparisons (Eliasziw et al., 1994). The mean square variances of some of the groups in this study were less than a fifth of other groups (see **Tables 8.7, 8.8, 8.11, 8.12, 8.17** and **8.18**) (Lahey, Downey & Saal, 1983; Mitchell, 1979).

The grip strength values of both hands of the males were significantly more reliable than those of the females, whether retested after one or four weeks. As the variances of the females were much less than that of the males, it is possible that the lower *ICC*s of the females were merely a reflection of their lower variances, as the *SEM*s, their 95% CI and their *MDE*s were quite similar to the Male LIG group. However the mean percentage change in the grip strength of both hands in the females was always greater than that of the males. Also the two-way ANOVA for gender and time confirmed the fact that the scores of the females were significantly less reliable than those of the males.

In these gender and time interval groups, the amount of change in the grip strength needed in either hand for that change to be considered to be beyond measurement error ranged from 4.53 kg f to 6.97 kg f. No other studies were located that examined the effect of short versus long retest time intervals on the reliability of grip strength values between the genders.

(c) The Reliability of the Grip Strength scores of the Dominant and Non-Dominant Hands, Split by Age and Gender

When sorted by age and gender, all groups had good to excellent *ICC* (3,1) values, ranging from .718 to .946. However, the grip strength scores of the 13-year-old females were significantly less reliable for both their hands than the 13-year-old males, due to no overlap in the *ICC* 95% *CI*s for either hand. The non-dominant handgrip strength scores of the 17-year-old females were also significantly less reliable than those of the 17-year-old males, with a marginal overlap in the 95% *CI*s for their dominant hand scores. To confirm this difference in grip strength reliabilities the ANOVAs demonstrated a weak but significant effect of gender for both hands.

Neither age, nor the interaction of age and gender, significantly affected the absolute mean percentage change in grip strength, which indicated that the lower *ICC*s for the females were again not solely due to their lower variances.

Some studies have examined the reliability of grip strength scores in teenagers, for either or both genders. Hinson et al. (1990) used a Jamar™ dynamometer and a one-day retest interval for 50 boys, aged 15-years-old. When using Pearson's Product Moment Correlation Coefficient (*r*) they found that the right hand achieved an $r = .954$ ($p < .001$) and the left hand $r = .946$ ($p = < .001$). These values were only slightly higher than the *ICC* (3,1) values of either of the current groups of males. As stated earlier, as the *ICC*s are a more conservative and appropriate measure of reliability than *r,* the collective current group of males have comparatively excellent reliability values for the scores of both hands.

Haward and Griffin (2002) found Spearman's *rho* reliability indices from the initial test to the first retest one week later of *rho* = .77 for the dominant hand and *rho* = .84 for the non-dominant hand in a group of 18 males, aged 18 to 25 years. The dominant and non-dominant hands of the current 13 and 17-year-old males all achieved higher reliability values (*ICC* = .910 to .946). Over a 10-week period Reddon et al. (1985) found that a sample of 6 right-handed women achieved slightly higher reliability coefficients than 6 men (.91 versus .94). The current results found the reverse to be true; the males had consistently higher reliability indices than the females.

Reliability studies with only females

Niebuhr et al., (1994) found that the *ICC* (3,1) values for a group of 28 young female adults (mean age = 24 years, $n = 33$) were slightly higher for the left hand than the right hand (.942 versus .930), although this difference was not significant. A group of 107 adult urban Australian women were repeat grip strength tested after 1 to 2 weeks (Hunter et al., 2000). These researchers obtained *ICC* (2,1) values of between .91 and .95 for a number of strength tests, including maximal voluntary grip strength with force transducers. Another Australian study (Nitschke, McMeekan, Burry & Matyas, 1999) found that the *r* value for grip strength retested over a 4 to 7 day period was $r = .93$ for 32 healthy women and $r = .95$ for 10 women with non-specific regional pain in their upper limb. Mathiowetz et al. (1984) found that the grip strength scores of the left hand ($r = .915$) had higher reliability values than the grip strength scores of the right hand ($r = .822$) in 27 Occupational Therapy students for the highest score of three trials repeat tested after seven days or less. Bohannon and Schaubert (2005) and Schaubert and Bohannon (2005) also found that the left hand had slightly higher *ICC* (3,1) values for grip strength scores in elderly people; each measurement being over .91 with the left grip strength reliability being .04 greater than that of the right. Thus women beyond the teenage years have achieved higher reliability indices than the present group of female teenagers whose *ICC* (3,1) values were not above .76 when divided by age, or retest time interval. This reliability figure improved for some when they were grouped by handedness classifications, as discussed in the next section.

For the females, the amount of change in the grip strength needed in either hand for that change to be considered to be beyond measurement error ranged from 4.63 kg f to 6.65 kg f. Extensive database searches failed to locate any research with which to compare these specific results.

(d) Grip Strength Reliability of the Hands Based on Degrees of Handedness

The grip strength reliabilities for the right-handed and mixed-handed teenagers were excellent and all equal to or above *ICC* (3,1) values of .959. In stark contrast, the scores of both hands of the left-handed teenagers demonstrated poor reliability with *ICC* values between .356 and .472. The variance values, especially for their non-dominant hand, were low. This may have adversely biased their results. The results of the left-handed teenagers were double-checked for translation errors; there were none. No comments were recorded at the time of the assessment that would explain the poor results for these teenagers (eg., sickness, injury, etc.). Two of these teenagers had grip strength changes of 21.4% and 30.0% in their dominant hand. Another two had grip strength changes of 20.0% and 22.0% in their non-dominant hand. With only a sample of 10 left-handed teenagers these hands represented 20% of their hands. In contrast, only two of the right-handed teenagers (2.3% of their group, n = 86) had dominant hand grip strength changes of 20% or greater (one with 20% and one with a 24% change) and three of these right-handed teenagers had non-dominant grip strength percentage changes between 20 and 21%. There were six mixed-handers (n = 58) who had dominant handgrip strength changes equal to, or greater than 20%; four of their non-dominant hands were in this category also. It would be unreasonable to draw any conclusions from the small sample of left-handed teenagers. A much larger sample of left-handers may raise the *ICC* values, based on the assumption that the current sample happened by chance, to have two such individuals who could not generate more consistent grip strength readings over time.

Within the right and mixed-handedness groups, the *ICC* values of the females were significantly lower than that of the males, for both dominant and non-dominant hands. For same-gender groups, there were no significant differences in the *ICC* values of either hand for either handedness group. Although all groups had similar *SEM*s and *MDE*s, see **Tables 8.17** and **8.18**.

The ANOVAs for the dominant and non-dominant hands of each gender and handedness group demonstrated only a weak but significant effect of gender on absolute mean percentage change in grip strength between the right- and the mixed-handed teenagers. This indicated that the lower *ICC*s for the females were again not solely due to their lower variances. In these right- and mixed-handed groups the amount of change in the grip strength needed in either the dominant or non-dominant hand for that change to be considered to be beyond measurement error ranged from 4.96 kg force to 6.28 kg f, depending on the gender, hand and handedness group of the teenager in question.

LIMITATIONS OF THE EXPERIMENT

A few researchers have retested adult participants more than twice. Niebuhr et al. (1994) and Haward and Griffins (2002) tested their young adults three times and Reddon et al. (1985) tested a sample of 12 adults ten times. The latter two research groups used weekly retest intervals. The current sample could have been repeat-tested three or more times to study their reliability patterns also. Haward and Griffins found that when assessed by Friedman's two-way analysis of variance by ranks their healthy young males had no significant difference in grip strength performance between three successive sessions for either the dominant or non-dominant hands.

In contrast, Niebuhr et al. found that the peak force reduced over the three sessions, but the reduction was only significant between the first and second sessions, which were spaced four weeks apart. The second and third sessions were spaced two weeks apart, due to timetable commitments of the participants. The left hands were significantly weaker, but their scores were not significantly more reliable than that of the right hands, as found in the female teenagers of the current study. Reddon et al. did not state which reliability coefficient they used, but found that:

> "The reliability for men was .91 and ranged between .81 and .99 and .79 to .99, for preferred and nonpreferred hands, respectively. For women the average test-retest reliability was .94 and ranged between .83 to 1.0 and .78 to .99, for preferred and nonpreferred hands, respectively." (1985, p. 1196)

As the results from the current experiment are comparable with the ranges found by these other studies, more repetitions of the grip strength testing sessions may not have provided increased knowledge about the grip strength behaviour of teenage males, but may have yielded more information about teenage females. The other limitation in this study was the lack of left-handed teenagers, because although their data appears to generate low reliability values, larger numbers of left-handed teenagers would be needed to confirm or reject this finding. Researchers have needed to specifically advertise for left-handed participant (Provins & Magliaro, 1993), test entire populations, or use twin registers to boost their numbers of left-handed participants.

SUMMARY

Clinicians need to know that retesting the maximal voluntary isometric grip strength testing of teenagers over either a one or four-week period produces excellent levels of reliability, with *ICC* (3,1) values ranging from .90 to .97 and there are no significant differences between the reliability values of the dominant and non-dominant hands. The inherent variability and errors of measurement mean that the clinician using grip strength testing to measure improvement must look for minimum changes in the grip strength readings that can be considered as real changes. In this study these were between 4.5 and 7.6 kg force. However, within these overall excellent results hide some teenagers with poor reliability for their grip strength readings. This analysis has detected them.

Males achieve significantly higher reliability values for the grip strength measurements of both their hands (*ICC*s from .94 to .99) than females (*ICC*s from .72 to .81), when tested over either a one or four-week period. The *MDE*s are similar for both genders over these time periods, and range from 4.5 to 7.0 kg f. But the absolute percentage change for the females can be between 30 to 45% greater than the males at 8.0 to 9.6 kg f, as compared to 5.3 to 6.5 kg f for the males.

For both their hands, thirteen year-old females have significantly lower reliability values for their grip strength readings than 13-year-old males. At 17 years of age, the grip strength reliability of the dominant hand of females is not significantly different to that of the dominant hand of 17-year-old males, even though their reliability values are notably lower and the grip strength readings of their non-dominant hands are significantly less reliable than the males. Slightly larger *MDE*s are displayed by both genders of 17-year-olds (6.0 to 7.0 kg f) than both genders of 13-year-olds (4.6 to 5.9 kg f).

The most interesting results were obtained when the teenagers were divided into three handedness groups. These comprised of right-, mixed- and left-handed teenagers. When the results for the two genders were pooled there were excellent reliability values for the right and mixed-handed teenagers, all above .96. There was no significant difference between the retest values of either hand, in or between either

group, with the *MDE*s ranging from 5.6 to 6.2 kg f. The grip strength readings for both hands of the left-handers had significantly and notably lower *ICC* values than the grip strength readings of either hand in the other two groups. The 95% *CI*s for the *ICC*s of the left-handers ranged into negative figures. After this group was deleted from further calculations the right and mixed-handed teenagers were divided by their gender. Within the genders there were no significant differences between the reliability values for either hand within either handedness group, or between the handedness groups. But between the genders, the *ICC* values for the males were always significantly more reliable than those for the females.

Thus age, gender and handedness do have significant effects in grip strength reliability. Left-handed teenagers appear to have very low reliability values for their grip strength readings. But as the variance in grip strength readings was very low for these left-handers it may have adversely biased their results, thus no firm conclusions should be drawn from their results.

Possible reasons for the lower reliability values for the grip strength scores of the females were that they perceived higher pain, or discomfort when squeezing the handles, or that they were inexperienced at repeatedly giving maximal efforts in tests of strength.

It appears that this is the first time that the reliability of the grip strength measurements of teenagers has been assessed when considering degrees of handedness and clinically relevant time intervals of one or four weeks.
This sample of teenagers was representative of urban Australian teenagers, as shown in the previous chapters, thus these reliability results can be generalised to the wider population of Australian teenagers. These reliability results are consistent with the general trends in other studies based on children and adults.

Further experiments concerning the reliability of grip strength measurements with female teenagers and with left-handed teenagers are needed. Physiological differences in muscles between the two genders (discussed in Chapter 3), may aid in accounting for the grip strength reliability differences of the two genders seen in this study, but they will not account for the handedness effect.

CHAPTER 9

THE RELIABILITY OF THE ISOMETRIC GRIP STRENGTH RATIO OF 154 TEENAGERS

It was established in Chapter 8 that the absolute mean percentage change in grip strength for each hand was between 4.85 and 6.97% for the males and 7.91 and 9.93% for the females, depending on whether they were divided by age, or handedness groups. However it has not been established whether the grip strength of the two hands increases or decreases in synchrony with each other, from an initial test session (T1) to a second test session (T2). Thus it is not known whether the two hands can validly act as reference points for each other, with changes in the strength of one hand reflecting changes in the other hand. In this chapter the reliability of the grip strength ratio (GSR) of the two hands will be examined. This GSR has been defined in several ways, but the most common method is to divide the maximal voluntary grip strength (grip strength) reading of one hand by that of the other hand.

When improvement in a hand injury has stopped, it is common for the bilateral grip strength of the patient to be measured for the purposes of writing a discharge summary, or a medico-legal report. If this testing is repeated several weeks later as part of an independent medico-legal assessment, then it should be expected that the GSR between the two hands would remain the same, although the absolute grip strength for each individual hand fluctuates from day to day in the normal population. The question arises as to what is the normal variation for the GSR, as it may be argued by the defendant in a medico-legal trial, that variation above a certain level is evidence of malingering. Alternatively the plaintiff may argue that such variation is normal or due to a lack of full recovery.

Thus the research question arises, should the GSR between the two *uninjured* hands be expected to be constant over time periods of several weeks? Periods between one to four weeks are common times between strength evaluations (depending on the diagnosis and the stage of the healing process). If the GSR in the two *uninjured* hands does fluctuate over such time periods, it is important to know the magnitude and direction of the change and to be able to identify the amount of change in the GSR.

Recently, the strength difference between the two hands has been expressed in terms of either the "grip strength ratio", or the "percent difference" between the dominant and non-dominant hands (Armstrong & Oldham, 1999; De Smet & Vercammen, 2001; Emerson et al., 2001; Häger-Ross & Rösblad, 2002; Incel et al., 2002; MacDermid et al., 2002; Massey-Westropp et al., 2004). In older studies it has been expressed as the difference between the major and minor hands (Schmidt & Toews, 1970), or the right and left hands (Fike & Rousseau, 1982). As discussed earlier, the methods for assessing handedness have varied and the equipment used for measuring grip strength has also not been consistent. However, the current consensus is that the right hand of right-handed people tends to be up to 10% stronger than the left hand, but that the left hand of left-handed people is often equal to the strength of the right hand (Bohannon, 2003). Generalisations can be unfair to the individual due to the diversity of grip strength differences between the two hands, as presented in Chapter 4 and in Clerke and Clerke (2001).

Here the GSR has been presented as the strongest grip strength value of the dominant hand divided by the strongest grip strength value of the non-dominant hand. Thus if a person had a GSR of 1.10 this would indicate that their dominant hand was 10% stronger than their non-dominant hand and conversely a GSR of 0.87 should be interpreted as the dominant hand having 87% of the strength of the non-dominant hand. Using this notion and the above general consensus, the average GSR of right-handed and left-handed people is 1.10 and 1.0 respectively.

In a study by Fike and Rousseau (1982), the right/left GSR, as recorded with a Jamar™ handgrip dynamometer, was found to be 1.07 for males and 1.10 for females; whereas in the same study when a Vigorimeter™ was used, the right/left GSR was recorded as 1.02 for males and 1.04 for females. This change in the GSR may be attributed to the fact that the first tool measured isometric (static) force and the latter measured dynamic grip strength, or pressure, indicating that the performance of the two hands in relation to each other can change, depending on the chosen measurement tool and the muscular demands required to operate it.

Only two studies were found in which the authors divided their participants into more than dichotomous handedness groups (Yim et al., 2003; Zverev & Kamadyaapa, 2001). When studying 102 males and 74 females between the ages of 15 and 66, Zverev and Kamadyaapa found that the mean difference between the maximal isometric grip strength (MGS) of the two hands varied by 10 to 20%. "For most of the left and right handers, the MGS of the preferred hand was highest while most of the mixed handed subjects exhibited higher MGS of the left hand" (p. 613). These authors presented a summary table of mean grip strength values for the left and right hands of the three handedness groups. When the values for all the right hands were divided by the values for all the left hands the mean right/left ratio was obtained. This resulted in the mean group GSR of the right-handers (n = 158) being 1.03, the left-handers (n = 10) being 1.06 and the mixed-handers (n = 8) being 0.97. They also added that "about half of the differences in MGS values of the two hands in [the] overall sample of males [was] less than 10%. In females, 30 to 40% of the relative differences between hands fell within [the] 10% category." (p. 613) The GSR of the individual can be quite different to that of the group mean.

When Yim et al. (2003) studied 712 children aged between 7 and 12, they divided the children into those who held a pencil and threw a ball with their right hand (defined as right-handed), or with their left hand (defined as left-handed), or with both hands ("both-hand-users"). From reading their grip strength graphs, it appears that the greatest right/left differences for the 12-year-old males (n = 38) was for those who were right-handed, with the GSR being estimated at 1.16. The GSR of the left and "both-hand-user" males appeared to be close to 1.0. Consistent with the males, the 12-year-old females (n = 41) appeared to have a GSR of 1.08 for the right-handers and 1.0 for both the left-handers and "both-hand-users". This GSR was not consistent over the different age groups from 7 to 12-year-olds and there did not appear to be any pattern of right/left strength differences throughout the age groups.

Prediction equations have been developed for the strength of one hand, based on the strength of the other hand, and have included other factors such as hand dominance (Crosby & Wehbé, 1994; Desrosiers et al., 1995; Hanten et al., 1999). These equations were not able to fully predict the grip strength of the hand in question. The unexplained variability present in such equations may be in part due to GSR fluctuations. This would occur when fluctuations in the grip strength of the two hands were in the opposite direction simultaneously. This would occur when one hand had an increased grip strength, whilst the other hand had a decreased in grip strength, compared to a previous test session. No studies on the degree of reliability of the GSR using Jamar™-like handgrip dynamometers were found. However, studies were found that dealt with the reliability of the grip strength scores of the individual hands over time (as presented in Chapter 8). There is a need to examine the issue of reliability of the GSR using these tools, due to their continuing application in clinical and research settings (Rajan et al., 2005; Shechtman, Gestewitz, et al., 2005; Shechtman, Gutierrez, et al., 2005).

From these considerations two null hypotheses have been proposed:

1. The amount of change in the GSR from the first test (T1) to the second test (T2) would not be affected by the independent variables of gender, retest time interval, or the interaction of these two independent variables;

2. The amount of change in the GSR from T1 to T2 would not be affected by the independent variables of gender and age, or the interaction of these two independent variables.

The following analyses of the grip strength data were carried out to: (a) determine the reliability of the GSR, and (b) to determine whether the three variables of retest time interval, age and gender affected the reliability of the GSR.

METHODS

The data used for this chapter were taken from the repeat-tested junior (13-year-old) and senior (17-year-old) teenagers studied in Chapter 8.The GSR was calculated from the maximal isometric grip strength of the dominant hand divided by the maximal isometric grip strength of the non-dominant hand in the initial test session. This was labelled GSR T1. The GSR of the second session was labelled GSR T2.

STATISTICAL ANALYSIS

For each group the means of the grip strength ratio (GSR) was determined by first calculating the GSR for each individual, and then averaging them (Newman et al., 1984; Petersen et al., 1989). Such ratios are termed mean individual ratios; this is to differentiate them from mean group ratios. Standard deviations for the GSRs were calculated for the teenagers after being sorted by gender and time interval groups, and then again after being sorted by gender and age groups. To calculate the reliability, screen for systematic errors, and to determine the amount of change in the GSR that would be considered to be beyond mere fluctuations, the following mathematical procedures were applied to each teenage group for their GSR:

1. Intraclass correlation coefficient model *ICC* (3,1) (Shrout & Fleiss, 1979) and their level of significance as determined by the *p* value derived from the *ICC* process in SPSS Version 11;
2. Percentage Close Agreement (PCA) (Rey et al., 1987);
3. Mean of the differences and mean of the absolute differences between T1 and T2;
4. Percentage coefficient of variation of the method error (%CV of ME), which equals 2ME divided by the sum of the means of T1 and T2, times one hundred. ME equals the standard deviation of the differences divided by the square root of two (Portney & Watkins, 2000);
5. Standard Error of Measurement (*SEM*), as described in Chapter 5;
6. Two-sided 95% confidence intervals (95% *CI*) for the *SEM* (Armitage & Berry, 1994);
7. Variance of the scores, as estimated from the total mean squares of the univariate ANOVAs. The homogeneity of the students within each group was also taken from these same ANOVAs, based on the subjects' mean square (SMS) estimates;
8. Two-way independent groups ANOVAs for pairs of independent variables and any interaction effect between these two;
9. Scatter plots of the relationship between the percentage change in the GSRs and the independent variables.

Statistical procedures numbered 1 and 2 were needed to examine the reliability of the ratios. *ICC* (3,1) was the appropriate reliability coefficient to use, as only one rater was the rater of interest, this being the sole tester, the author. The second procedure, the PCA values were made by calculating the GSR T1 scores/GSR T2 scores, then multiplying by 100 (Moseley & Adams, 1991; Rey et al., 1987). This current experiment grouped the PCA values in 5.0% point divisions. The PCA is one way of expressing the change in the GSR from the initial to the retest by simply stating the amount of change that occurred without taking into account chance agreement. It will be an overestimation of true reliability, but is a quickly calculated and easily understood procedure for clinicians and patients (Portney & Watkins, 2000). The third set of procedures was needed to express the amount of change in the ratio scores from the T1 to T2, with and without direction.

Portney and Watkins (2000) point out that when there is a range restriction, as there was with the GSRs of the female groups, a more appropriate statistic to express the amount of error to expect as measurement error is the %CV of the ME, as opposed to solely focusing on reliability indices. Thus the fourth procedure was included. Procedures numbered 5 and 6 examined the difference in the ratios that would be needed to regard a change in the ratios as a real change. Procedure 7, the variance of the scores, was added to assess whether the *ICC*s could validly be compared across the two genders. This is because groups with low variances can indicate a restriction in the range of values that may be the cause of low *ICC*s. Such groups should not be compared with groups that have much higher variances (Portney & Watkins, 2000).

Procedure 8 was the two-way independent ANOVA procedure for testing the two null hypotheses concerning the effects of the independent variables of gender, retest interval duration and age on the amount of change in the GSR. Each independent variable had two levels, with retest interval duration designated as either short (SIG) or long (LIG). A 5% level of significance was chosen. The dependent variable was the difference between GSR T1 and GSR T2 for each teenager. The strengths of any significant findings were tested using eta squared. Scatter plots (procedure 9) were constructed to visualise the differences against the mean of the tests, to help identify the presence of outliers and the normality of the differences (Bland & Altman, 1986; Portney & Watkins, 2000).

RESULTS

1. Reliability of the GSR, Sorted by Gender and Retest Time Interval

The mean individual ratios for the initial and the repeat tests as derived from the four student groups; Male Short Interval Group (SIG), Male Long Interval Group (LIG), Female SIG and Female LIG were all similar, but there was a tendency for the mean to decrease on T2, whilst the *SDs* tended to widen (see **Table 9.1)**. The test-retest differences and their results for the GSR of these four groups as assessed by the various statistical procedures numbered 1 and 3 to 7 are presented in **Table 9.2**.

Table 9.1. Grip Strength Ratios (GSR) for First (T1) and Second Test (T2) Sorted by Gender and Time Interval.

Group	GSR T1 (*SD*)	GSR T2 (*SD*)
Male SIG[a]	1.08 (0.10)	1.07 (0.12)
Male LIG[b]	1.08 (0.09)	1.07 (0.09)
Female SIG[c]	1.11 (0.11)	1.11 (0.14)
Female LIG[d]	1.11 (0.11)	1.07 (0.11)

Note. SIG = short retest interval group; LIG = long retest interval group.
[a]*n* = 40. [b]*n* = 36. [c]*n* = 44. [d]*n* = 34

The two-way independent groups ANOVA in **Table 9.3** showed no significant main effects of time interval, or gender, or any interaction between the two.

The PCAs for the gender and time interval groups have been presented in the bar graph of **Figure 9.1**. All groups except the Female LIGs had similar percentages of teenagers who changed their GSR by 5, 10, 15 and 20%. There was a large percentage of Female LIGs (17.4%) who had a change of 15.1 to 20% in their GSR. When Percentage Close Agreements (PCA) were analysed incorporating the direction of the change in the GSRs, the Female LIGs once again exhibited a different pattern to the others, namely their larger percentage of participants (59%) who increased their GSR. For the Male SIGs there were 17 (42.5%) who decreased their GSR and 21 (52.5%) who increased their GSR and two (5%) who did not change their GSR. For the Male LIGs there were 18 (50%) who decreased their GSR and 15 (41.7%) who increased their GSR, and three (8.3%) who did not change their GSR. For the Female SIGs there were 21 (47.7%) who decreased their GSR and 21 (47.7%) who increased their GSR, and two (4.5%) who did not change their GSR. For the Female LIGs there were 12 (35.3%) who decreased their GSR and 20 (58.8%) who increased their GSR, and two (5.9%) who did not change their GSR.

Table 9.2. Statistical Analysis of the GSR, Sorted by Gender and Time Interval.

	Male		Female	
Statistic	SIG[a]	LIG[b]	SIG[c]	LIG[d]
ICC (3,1)	.652	.315	.635	.497
ICC 95% *CI*	.431 – .800	-.011 – .580	.419 – .783	.196 – .712
Mean Diff	0.009	0.006	0.001	0.044
Mean A Diff	0.069	0.08	0.081	0.094
%CV of ME	6.78	7.02	6.71	15.09
SEM	0.064	0.075	0.076	0.078
SEM 95% *CI*	0.053 – 0.083	0.061 – 0.098	0.063 – 0.096	0.061 – 0.100
MDE	0.177	0.207	0.211	0.215
Variance	0.012	0.008	0.016	0.012

Note. SIG = short retest interval group; LIG = long retest interval group; Mean A Diff = mean of the absolute difference.

[a]$n = 40$. [b]$n = 36$. [c]$n = 44$. [d]$n = 34$

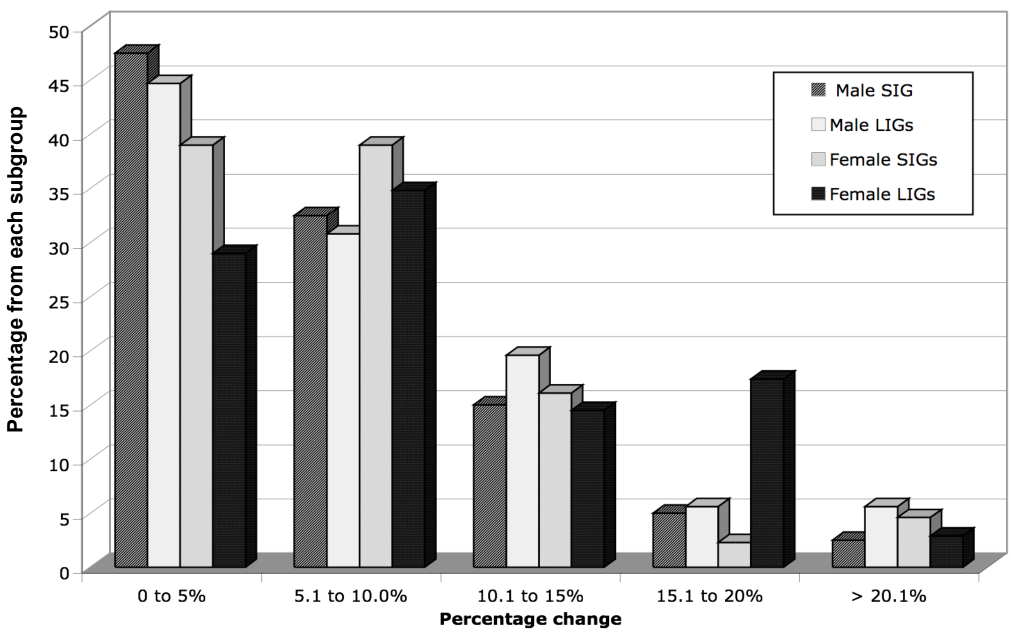
$p < .01$ for all groups, except Male LIG with $p = .03$.

The *ICC* (3,1) values demonstrated that the scores of the Male SIGs achieved the highest reliability and that of the Male LIGs achieved the lowest, see **Table 9.2**. The mean of the differences from T1 to T2 were less than 0.01 for all groups. The mean of the absolute differences revealed a maximum of nearly 0.01 for the Female LIGs that would be a change of approximately 10% in the GSR fromT1 to T2.

Table 9.3. Two-Way Independent Groups ANOVA for Gender and Time Interval for the GSR.

Source	Sum of Squares	*df*	Mean Square	*F*	*p*
Gender	58.08	1	58.08	1.388	.241
Time Interval	51.53	1	51.53	1.232	.269
Gender x Time	21.95	1	21.95	0.00	.982
Error	6276.33	150	41.84		
Corrected Total	6382.25	153			

Figure 9.1. Percentage close agreement for GSR sorted by gender and time interval.



2. Reliability of the GSR, Sorted by Gender and Age

The mean individual GSRs derived from the groups sorted by gender and age were close, varying from 1.08 to 1.13 for T1 to T2 with *SD*s ranging from 0.08 to 0.14, see **Table 9.4**. As seen with the gender and time interval groups, there was again a tendency for the mean GSRs to decrease at T2, but for the *SD* to widen. The test-retest differences and results for these groups, as assessed by the various statistical procedures numbered 1 and 3 to 7 are presented in **Table 9.5**.

Table 9.4. Gender and Age Split GSR for T1 and T2.

Group	GSR T1 (*SD*)	GSR T2 (*SD*)
13 y.o. males[a]	1.07 (0.11)	1.06 (0.10)
17 y.o. males[a]	1.09 (0.08)	1.08 (0.12)
13 y.o. females[b]	1.08 (0.09)	1.08 (0.11)
17 y.o. females[c]	1.13 (0.13)	1.10 (0.14)

[a]n = 38. [b]n = 34. [c]n = 44.

The two-way independent groups ANOVA in **Table 9.6** displayed no significant main effects of age, or gender, or any interaction between the two.

The PCAs for the four age and gender groups have been presented in the following bar graph, **Figure 9.2**. The 13-year-old males had the highest percentage of participants (47.5%) who changed their PCA by less than 5%, whilst the 17-year-old females had the greatest percentage of participants (11.4%) with a 15.1 to 20% change in GSR. However all groups had a similar percentage of participants who changed by 10% or less. Thus there were 28 (73.7%), 26 (76.5%), 28 (73.3%) and 30 (68.2%) 13-year-old males and females, 17-year-old males and females respectively, who changed by less than 10%.

Figure 9.2. Percentage close agreement for GSR sorted by age and gender.

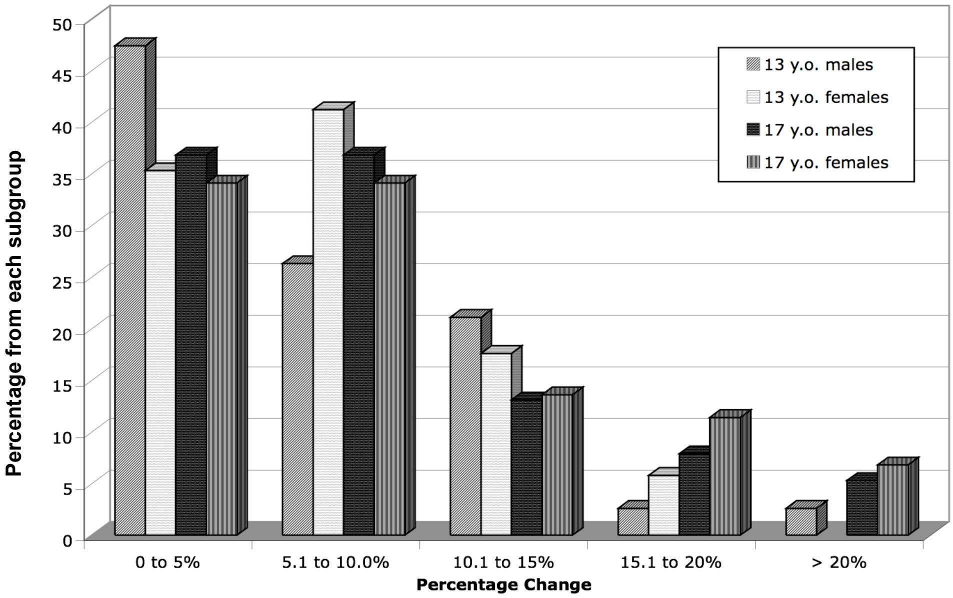


Table 9.5. Statistical Analysis of the GSR, Sorted by Gender and Age.

	Males[a]		Females[b]	
Statistic	13 y.o.	17 y.o.	13 y.o.	17 y.o.
ICC (3,1)	.624	.407	.558	.576
ICC 95% *CI*	.384 – .785	.104 – .640	.275 – .752	.340 – .744
Mean Diff	0.007	0.008	0.013	0.025
Mean A Diff	0.07	0.083	0.076	0.095
%CV of ME	6.35	7.51	7.03	12.47
SEM	0.062	0.075	0.065	0.086
SEM 95% *CI*	0.051 – 0.082	0.062 – 0.098	0.053 – 0.086	0.071 – 0.108
MDE	0.173	0.208	0.181	0.238
Variance	0.01	0.01	0.01	0.017

Note. Mean diff. = mean of the difference; Mean A Diff = mean of the absolute difference.
[a]n = 119. [b]n = 116.
$p < .01$.

Table 9.6. Two-Way Independent Groups ANOVA for Gender and Age for the GSR.

Source	Sum of Squares	*df*	Mean Square	*F*	*p*
Gender	43.086	1	43.086	1.039	.310
Age	106.982	1	106.982	2.581	.110
Gender x Age	3.199	1	3.199	0.077	.782
Error	6217.550	150	41.450		
Corrected Total	6382.252	153			

When the PCAs were analysed incorporating the direction of change, the 13 and 17-year-old females had slightly more participants with an increase in their GSR, whilst the 17-year-old males had the largest percentage of participants who decreased their PCAs. For the 13-year-old males, there were 15 (39.5%) who decreased their GSR, 18 (47.3%) who increased their GSR and 5 (13.2%) who did not change theirs. For the 13-year-old females, there were 14 (41.2%) who decreased their GSR, 18 (52.9%) who increased their GSR and 2 (5.9%) who did not change theirs. For the 17-year-old males, there were 20 (52.6%) who decreased their GSR, 18 (47.4%) who increased their GSR, but no one in this group had an unchanged GSR. For the 17-year-old females, there were 19 (43.2%) who decreased their GSR, 23 (52.2%) who increased their GSR and 2 (4.5%) who remained unchanged. Overall there were 112 (72.7%) teenagers who had a GSR change equal to or less than 10% from T1 to T2, with 6 (3.8%) who had a change greater than 20%.

The *ICC* (3,1) values demonstrated that the GSR of the 13-year-old males had the highest reliability and the 17-year-old males had the lowest, see **Table 9.5**. The mean of the differences from T1 to T2 were close to zero, being about 0.01 for all groups, except the 17-year-old females, who had a value of 0.025. The mean of the absolute differences reached a maximum of nearly 0.01 for the 17-year-old females. This represented a change of about 10% in the GSR from T1 to T2. The *SEM* values indicated small changes within and between groups. The 95% *CI*s for the *SEM* values were fairly uniform in width, but had lower limits between 0.051 and 0.071 with upper limits between 0.082 and 0.108.

DISCUSSION

In Chapter 8 the reliability of the grip strength measurements for the two hands was examined separately. The measurements of the dominant hand achieved higher reliability values than those of the non-dominant hand. Knowing this, it would largely be the change in the grip strength of the non-dominant hand from test 1 (T1) to test 2 (T2) that altered the grip strength ratio (GSR) between the two test sessions. The amount of change required in the GSR to have the change considered to be a real change has been determined for each age and gender-split group, and each age and time interval group. This ranged from a GSR change of 0.173 (for the 13-year-old males) to 0.238 (for the 17-year-old females). The null hypothesis that the amount of change in the GSR from T1 to T2 would not be affected by the independent variables of gender, retest time interval, or the interaction of these two independent variables was found to be consistent with the data. Likewise the null hypothesis that the amount of change in the GSR from T1 to T2 would not be affected by age and gender, or the interaction of these two independent variables was similarly retained.

Gender and Time Interval

When the gender and time split groups were examined there was a trend for the GSR to decrease over time, whilst the *SD*s became wider for the male and female SIGs, as can be seen in **Table 9.1**. As judged by the *ICC* (3,1) values, the reliability of the GSR was highest for the Male Short Interval Group (SIG) (*ICC* = .652) and lowest for the Male Long Interval Group (LIG) (*ICC* = .315), with its lower 95% confidence limit having a negative value (-.011). Based on the cut offs suggested by Portney and Watkins (2000), these results can be interpreted as the GSR of the Male SIGs achieving good reliability and that of the Male LIGs achieving poor reliability, with that of the two female groups having fair reliability, except that the Female LIG group had a wider 95% *CI*. The *ICC* value of .652 for the Male SIGs means that only 65.2% of their variance in the GSR from T1 to T2 could be accounted for, leaving 37.5% of their variance unexplained. Portney and Watkins consider that *ICC* retest values for clinical measurements, such as grip strength, "should exceed .90 to ensure reasonable validity" (p. 565). Based on the current *ICC*s, none of these results meet this criterion. However, when the range of measurement values is quite restricted, *ICC* values can give an inaccurate representation. In such cases the %CV of ME is favoured (Portney & Watkins, 2000).

This inaccurate representation was seen when the statistical results of *ICC*s, Percentage Close Agreement (PCA) and %CV of ME were compared to express the reliability and amount of change in the GSR from T1 to T2. When the *ICC*s were considered the least reliable result was from the Male LIGs, but the Female LIGs had the greatest mean differences and greatest absolute differences in their GSR over time, with the greatest change according to the PCA figures and the highest %CV of ME (at 15.09%). The *ICC*s may have been affected by a restriction in range as the variances in these groups ranged from 0.008 to 0.016.

The %CV of ME can be considered to be a more rigorous method for examining change than the PCA, because unlike the PCA, it takes into account all types of error (Portney & Watkins, 2000). By using the %CV of ME it could be clarified that there was greater variability in the Female LIGs than in the other groups. The *SEM* and its 95% *CI* demonstrated that the Female LIG again had the highest amount of error and the widest confidence intervals. Using the *MDE* to express the magnitude of change that would be needed to be greater than normal fluctuation (as expressed it in the original units of measurement), the *MDE*s for the four groups were between a GSR change of 0.177 to 0.215. The group with the smallest *MDE* was the Male SIG group, then the Male LIG group, then the Female SIG, with the Female LIG groups having the biggest *MDE* values.

Gender and Age

As for the groups split for gender and time, when the gender and age split subgroups were examined, there was a trend for the GSR to decrease in the second session, whilst the *SD*s became wider, as seen in **Table 9.4**. As judged by the *ICC* (3,1) values, the reliability of the GSR was highest for the 13-year-old males (*ICC* = .624), lowest for the 17-year-old males (*ICC* = .407), and the two female groups had *ICC* values between these two male groups. Thus the GSRs of the younger males had good reliability and that of the older males had only fair reliability. The GSRs of all the groups had similar and wide *ICC* 95% *CI*s. The GSR of the 17-year-old males had the widest *CI*s and their lower *CI* approached zero (0.104). However, as for the gender and age split groups, these results need to be viewed in the light of other statistical methodologies before conclusions can be made.

The *ICC*s and PCAs again were contradictory. The *ICC*s showed that the group with the least reliable GSR was the 17-year-old males, but the PCAs showed that the GSR of the 17-year-old females had the greatest number of changes equal to, or greater than 15%. Also these females had the greatest absolute differences in their GSR over time. Thus again the restricted range of the GSRs may have influenced the *ICC* values. In support of the findings of the PCAs, when the %CV of ME results were considered, the reliability was much greater in the results of the 17-year-old males, than in those of the 17-year-old females. The group with the least amount of measurement error was the 13-year-old males (%CV of ME = 6.35%), whilst the group with the greatest amount of measurement error was the 17-year-old females (%CV of ME = 12.47%). The *SEM*, its 95% *CI* and its derived *MDE* confirmed that the 17-year-old females had the greatest amount of error and required the greatest amount of change in the GSR before the change could be considered to be a real change. The *MDE*s for the gender and age split groups were between 0.173 and 0.238; the measurements of the older teenagers required a greater change than that of the younger ones.

CONCLUSION AND FUTURE RECOMMENDATIONS

The relative reliability of the grip strength ratio (GSR) between the two genders, over two retest time intervals and between the two age groups of 13 and 17 years altered depending on the statistical test used. Due to the restriction in the range of the GSRs the most appropriate statistical method was the %CV of ME. Using this test, it would appear that the reliability of the GSR in the current sample of teenagers was relatively high for the males. If the females were tested after one week, or if they were 13 years old their measurements also displayed similar levels of good reliability. All these teenagers displayed a small %CV of ME in their GSR from T1 to T2 of up to 6 or 7%. However, if they were females and were retested after four weeks, or if they were 17 years old, the %CV of ME was as great as 12 or 15%. The ANOVAs found that neither gender, retest time interval, nor age had significant individual, or combined effects on the GSR from T1 to T2.

Another way to view these results is to observe that the magnitude of change needed in the GSR from T1 to T2 for it to be considered to be a real change, was 0.173 to 0.238. This depended upon which gender, time interval and age group the teenager belonged to. The values for the females represented relatively large amounts of error. Thus further research could be done with recovered and uninjured hands over multiple test sessions, spaced over clinically relevant time frames, of several weeks to several months, to ascertain whether the results for the females were normal and not a chance extreme result. Thus a proportion of the unexplained variance in the prediction equations of Chapter 7, Part 3 appear to be due to the grip strength values of the two hands fluctuating out of synchrony with each other.

Despite the above stated limitations it would appear that the GSR is an acceptably stable value for the majority of healthy teenagers. Hence clinicians should consider calculating the GSRs for their patients and then use them to help guide interpretations of observed changes in grip strength of one hand in relation to the other hand over time. However, clinicians should keep in mind that the GSR was not as stable for females as it was for males, especially if their results were compared over a four-week interval, rather than a one-week interval.

CHAPTER 10

HAND SHAPE EFFECTS ON MAXIMAL ISOMETRIC GRIP STRENGTH AND ITS RELIABILITY IN TEENAGERS

Most human anthropometric dimensions are distributed in a bell-shaped curve within the normal population, with slight male/female differences in the median and range values. These differences in relation to height, weight and hand dimensions have been explored in Chapter 7. Hand shape has been defined in various ways, but often as simply the hand width/hand length ratio (W/L ratio) (Kulaksiz & Gozil, 2002). Thus hands with varying W/L ratios can be described as 'long and narrow', 'average looking', or 'relatively square' by how long the hand is in relation to the width of the palm see **Appendix B**, **Figure 1**.

Paediatric studies have shown that the anthropometric variables of body size and hand length, palm length and palm width are highly correlated with grip strength in children. Males are stronger than females after the age of 12 or 13, and the dominant hand is generally stronger than the non-dominant hand, as discussed in Chapter 7. Also the maximal grip strength varies depending on which handle size is being used on the dynamometer (Firrell & Crain, 1996).

Researchers have shown that maximal grip strength varies throughout the day, and that there remains a large amount of this variation in the grip strength of healthy people that has not been accounted for, as stated in the preceding chapters. Specifically, the relationship between the shape of a hand and its ability to generate maximal and reliable grip strength readings with an isometric handgrip dynamometer has not yet been investigated. This chapter focuses on this aspect of grip strength force reading; also see **Appendix B** that is a copy of the published article concerning this topic.

Two hypotheses are examined in this chapter:

a) That there are differences in grip strength associated with hand shape, gender and hand dominance groups;

b) That there are differences in the reliability of grip strength measurements from an initial test (T1) to a repeat session (T2) associated with hand shape, gender and dominance.

METHODS

Details of the sample and inclusion criteria were presented in Chapter 5, the aims and methods for the thesis. The rationale for excluding outliers was given in Chapter 7. Thus the following analyses are based on the complete data sets of 232 teenagers (118 males and 114 females). The methods for measuring hand dimensions, grip strength testing methods and calibration of the GripTrack™ were presented in Chapter 5, with the results in the relevant sections of Chapter 7.

STATISTICAL ANALYSIS

Hand Shape, Gender and Maximal Grip Strength Force

Because the gender split frequency distributions of hand shapes were uni-modal, symmetrical and well approximated by normal distributions, partitioning at the quartiles formed three hand shape groups within each gender. This was done by using the frequency distributions of the W/L ratios. The relatively long-handed (labelled LONGhand) and relatively square-handed (labelled SQUAREhand) groups corresponded to those beyond the lower and upper quartiles, respectively. The average-handed (AVGEhand) group represented the 50% of hands between these two quartiles. Means, standard deviations and range for W/L ratio in these groups are given in **Table 10.1.**

Chi-square is a nonparametric test that was used to determine if there was a significant association between hand shape (considered to be nominal data) and maximum grip strength for the two genders. It was used to see if the observed frequencies of grip strength values within each hand shape group differed from the theoretical expected frequencies (Portney & Watkins, 2000). Along with descriptive statistics, the Chi-square analyses were undertaken using SPSS. An analysis of variance (ANOVA) with orthogonal contrasts (Winer, Brown & Michels, 1991) was used to determine the possible effects of, and interactions between, two between-groups factors (gender and hand shape) and one repeated measures factor (dominance, that it testing the dominant then the non-dominant hand). Planned trend contrasts examined for the presence of linear and quadratic trend components in possible strength changes across the three hand shapes. When a significant effect or interaction was found, the magnitude of the effect (the difference between the means of two compared groups) in the original units of measurement and their associated 95% confidence intervals (*CI*s) were calculated using the PSY software for contrast analysis (Bird, 2004). Contrast analyses are used to compare two means as part of a multiple comparison procedure to minimise Type 1 and Type II errors. Results were considered to be significant if their associated *p* values were less than .05.

Reliability of the Maximal Grip Strength Forces of the Dominant and Non-dominant Hands

The repeat test session was conducted with the same protocol as the original test session. The 149 teenagers who had their grip strength repeat tested were re-distributed into quartile groups, based on their W/L ratio. This resulted in slightly different quartile cut-off points compared to those made using the set of 232 teenagers. Six teenagers (four males and two females) shifted to an adjacent hand shape group via this process. The hand shape groups in this repeat-tested sample were labelled with the subscript $_2$ so that $LONG_2$, $AVGE_2$, and $SQUARE_2$ were distinguished from the previous larger samples. The reliability of the maximal grip strength results for dominant and non-dominant hands was then assessed from T1 to T2 using a single measure, the Intraclass Correlation Coefficient (*ICC* 3,1). The *ICC* (2,1) was also calculated as a check for the presence of additive biases between measures. To assess measurement accuracy for each of the 12 groups (2 gender x 2 dominance x 3 hand shape groups) the standard error of measurement (*SEM*) was calculated as $SEM = SD \times \sqrt{(1-ICC)}$ (Portney & Watkins, 2000). The standard deviation (*SD*) was the pooled standard deviation for T1 and T2, for each group separately. An ANOVA was performed to determine any systematic effects on grip strength. Differences between means and their *CI*s were calculated using the PSY software (Bird, 2004).

RESULTS

There were complete hand dimension data sets for 232 teenagers, with complete grip strength data sets for 228, once outliers had been excluded. The rationale for excluding outliers in this type of study was given in Chapter 7. Complete repeat grip strength data sets were available for 149 of these. Means, standard deviations and ranges for W/L ratio in each hand shape group and gender in the original sample are given in **Table 10.1**, and the W/L ratio means, standard deviations and ranges of the smaller, repeat-tested sample (hand shape groups $LONG_2$, $AVGE_2$ and $SQUARE_2$) are given in **Table 10.2**. Comparison of the cut off points for the hand shape groups between the two samples showed them to differ minimally, indicating a high degree of similarity between samples. Mean hand lengths, hand widths and their respective standard deviations have been reported for age and gender splits within the original sample of 232 and are presented in Chapter 7, Part 1a, **Table 7.3.**

Table 10.1. Mean Width/Length (W/L) Ratios, *SD*s and Ranges for the Original Male and Female Hand Shape Groups.

	Male		Female	
Group	W/L ratio Mean (*SD*)	Range	W/L ratio Mean (*SD*)	Range
LONGhand	0.517 (0.009)[a]	0.486 – 0.526	0.494 (0.010)[b]	0.472 – 0.508
AVGEhand	0.546 (0.012)[c]	0.527 - 0.564	0.530 (0.013)[c]	0.508 - 0.551
SQUAREhand	0.578 (0.019)[d]	0.565 - 0.665	0.567 (0.014)[e]	0.552 - 0.631

[a]*n* = 29. [b]*n* = 27. [c]*n* = 59. [d]*n* = 30. [e]*n* = 28.

Table 10.2. Mean Width/Length (W/L) Ratios, *SD*s and Ranges for the Repeat-Tested Sample of Male and Female Hand Shape Groups.

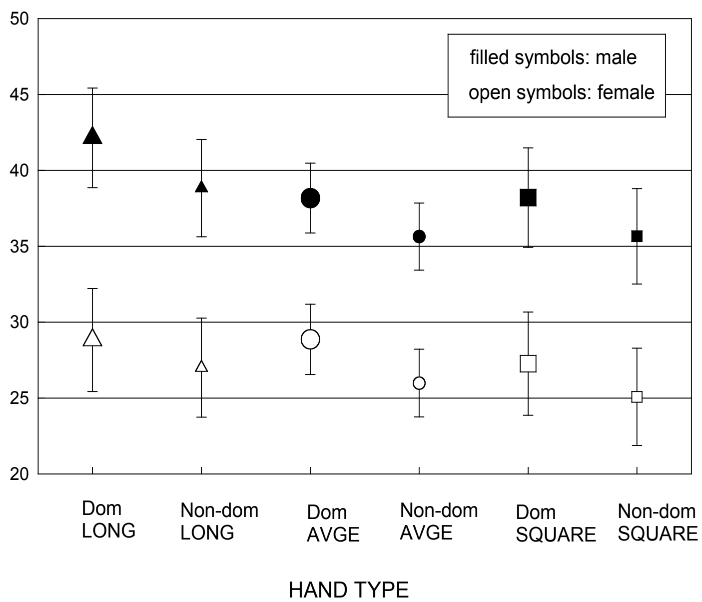
	Male		Female	
Group	W/L ratio Mean (SD)	Range	W/L ratio Mean (SD)	Range
$LONG_2$	0.520 (0.007)[a]	0.508 – 0.528	0.495 (0.010)[b]	0.472 – 0.509
$AVGE_2$	0.546 (0.010)[c]	0.529 - 0.565	0.527 (0.012)[c]	0.509 - 0.552
$SQUARE_2$	0.582 (0.023)[b]	0.566 - 0.665	0.572 (0.020)[b]	0.554 - 0.631

Note. $LONG_2$ = retested long hand group; $AVGE_2$ = retested average hand group; $SQUARE_2$ = retested square hand group
[a]n = 18. [b]n = 19. [c]n = 37.

Hand Shape, Gender and Maximal Grip Strength Force

Male and female grip strength force means and *CI*s for the dominant and non-dominant hands of each hand shape group are displayed in **Figure 10.1.**

Figure 10.1. Mean grip strength and 95% *CI* for males and females with six different hand types



Results from the ANOVA on the strength data (detailed in **Table 10.3**) from the original group showed males to be significantly stronger than females [$F_{(1, 222)}$ = 82.35, $p < .001$] with a difference between the means of 11.02 kg f (*CI*: 8.63 to 13.41 kg f). Dominant hands were found to be significantly stronger than non-dominant hands [$F_{(1, 222)}$ = 121.62, $p < .01$], giving a difference between the means of 2.53 kg f (*CI*: 2.08 to 2.98 kg f). The downward linear trend in strength from LONG, through AVGE to SQUARE hands, with the greatest difference between means being 2.96 kg f (*CI*: -0.27 to 6.18 kg f), was not statistically significant at the .05 level [$F_{(1, 222)}$ = 3.26, $p = .08$]. There was no significant quadratic trend in strength, nor any interaction effects between gender, dominance and hand shape.

Table 10.3. Male and Female Mean (*SD*) Grip Strengths Sorted by Dominance and Hand Shape.

		Mean kg f (*SD*)	
Gender and Hand Shape Group	*n*	Dominant Hand	Non-dominant Hand
Male LONGhand	28	42.32 (11.81)	39.64 (12.06)
Male AVGEhand	59	38.17 (11.36)	35.36 (10.37)
Male SQUAREhand	29	38.21 (11.71)	35.41 (11.82)
Female LONGhand	27	28.82 (5.62)	27.00 (4.65)
Female AVGEhand	58	28.86 (4.89)	25.98 (4.61)
Female SQUAREhand	27	27.26 (4.71)	25.07 (4.36)

Both genders and both hands showed similar patterns of optimal handle positions for generating maximal grip strength, the order of which was handle position (HP) 2 (75%), HP# 3 (19%), HP #1 (4%), HP #4 (2%) and HP#5 (0%). Chi square analysis showed no significant association between hand shape and handle position with respect to frequency of maximum grip strength, for either gender.

Reliability of the Maximal Grip Strength Forces of the Dominant and Non-dominant Hands

As with the larger group of teenagers, the males in the repeat-tested group were significantly stronger than the females (F (1,143) = 69.63, $p < .01$), with a difference of 12.38 kg f (*CI*: 9.48 to 15.31 kg f). Dominant hands of both genders were stronger than non-dominant hands (F (1, 143) = 87.95, $p < .01$), with a difference between the means of 2.41 kg f (*CI*: 1.90 to 2.91 kg f). Repeating the test yielded a small but significant increase in grip strength (F (1, 143) = 7.57, $p < .01$) with a mean increase of 0.62 kg f (*CI*: 0.17 to 1.06 kg f). There were no significant interactions involving gender, dominance, and hand shape.

For males, the grip strength reliability of both hands, in each hand shape group was excellent, as judged by the high *ICC* (3,1) values that ranged from .954 to .973. The *CI*s for these six *ICC* values overlap, indicating that the reliability of strength retest for male hand shape groups, whether using dominant or non-dominant hands, was not significantly different (**Figure 10.2**). For females, the test-retest *ICC*s became lower for both the dominant and non-dominant hands across the three hand-shape groups. Based on non-overlapping *CI*s, the $SQUARE_2$ groups had significantly lower *ICC* values than the most reliably measured group, the dominant $LONG_2$ group. The grip strength values of the non-dominant hand of the $LONG_2$ group also had excellent reliability; while the two $AVGE_2$ groups had moderately high reliability for repeat testing. Further, only female teenagers in the $LONG_2$ had strength retest reliability that was not significantly different from that of their male counterparts (**Figure 10.2**).

The calculated *ICC* (2,1) values were found to be smaller than the *ICC* (3,1) values throughout, but as this difference was in the second or third decimal place, these are not reported. The *SEM* is a measure of the magnitude of error associated with a test (Stratford & Goldsmith, 1997) and these values are presented in **Table 10.4**. Both males and females exhibited a similar range of *SEM*s for dominant and non-dominant hands, with the male dominant $LONG_2$ group having the greatest magnitude of errors. Grip strength values for the initial tests of the repeat-tested groups are presented in **Table 10.4.**

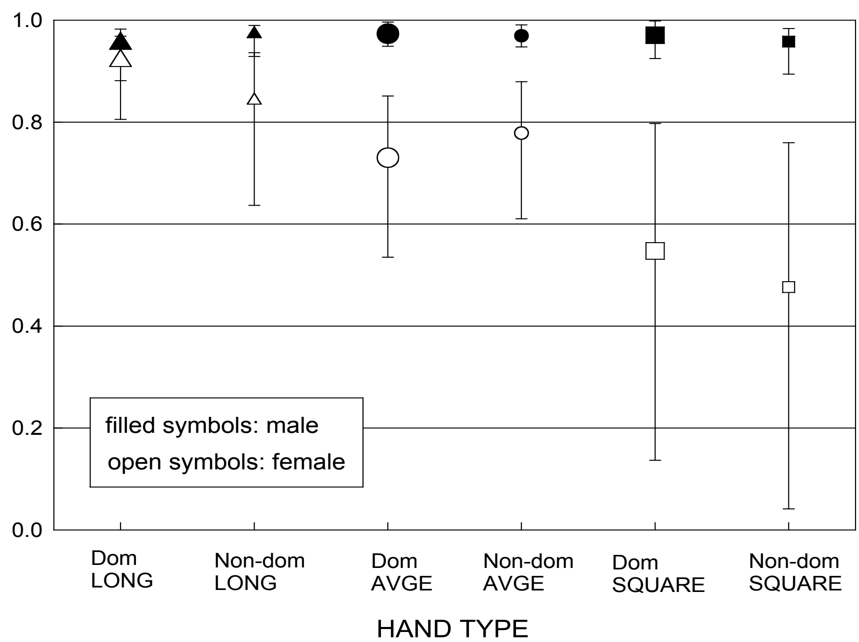
Table 10.4. Mean (*SD*) Grip Strengths (kg f) Sorted by Dominance and Hand Shape for Initial Test of Repeat-Tested Group With Standard Errors of Measurement (*SEM*) in kg f.

Gender and Hand Shape Group	*n*	Dominant Hand Mean (*SD*)	Dominant Hand *SEM*	Non-dominant Hand Mean (*SD*)	Non-dominant Hand *SEM*
Male $LONG_2$	18	44.11 (12.51)	2.75	41.33 (11.67)	1.95
Male $AVGE_2$	37	39.00 (12.00)	1.92	36.14 (10.42)	1.83
Male $SQUARE_2$	19	38.74 (12.05)	2.01	36.58 (12.44)	2.45
Female $LONG_2$	19	27.53 (5.71)	1.63	26.11 (5.11)	1.99
Female $AVGE_2$	37	28.87 (4.80)	2.38	25.76 (4.63)	2.10
Female $SQUARE_2$	19	27.90 (4.05)	2.58	24.79 (3.42)	2.49

Note. $LONG_2$ = retested long hand group; $AVGE_2$ = retested average hand group; $SQUARE_2$ = retested square hand group.

Figure 10.2. Reliability indices (*ICC*s) and 95% *CI*s for handgrip test-retest on males and females with six different hand types.

FIGURE 4. Reliability indices (ICCs) and 95% CIs for handgrip test-retest on males and females with six different hand types



DISCUSSION

Although the males were significantly stronger than the females, and the dominant hand was significantly stronger than the non-dominant hand, the shape of a hand did not significantly bias a teenager's grip strength force. But it can be noted that the long shaped hands tended to be stronger than the other two hand shapes.

There were substantial gender differences in grip strength reliability as a function of hand shape. The hypothesis that some hand shapes influenced the ability to generate a reliable grip strength force over time was not supported for the males, but was supported for the females. For the males both the dominant and non-dominant hands achieved excellent reliability values for grip strength irrespective of their hand shape category, all their *ICC* values were greater than .95.

In contrast, the grip strength readings of the females in the $SQUARE_2$ group showed poor reliability values (*ICC*= .547 and .476 for the dominant and non-dominant hands, respectively). The reliability values for this group were significantly poorer than the dominant and non-dominant hands of the $AVGE_2$, and the $LONG_2$ groups, whose *ICC* values ranged from .730 to .920 (**Figure 10.2**). There appears to be no other research with which to compare these findings.

The finding of handle position #2 being the most frequently optimal handle position, followed by handle position #3, was consistent with previous studies (Firrell & Crain, 1996).

The *ICC* is seen as representing the ability of a test to discriminate among participants, and the *SEM* as a measure of the magnitude of test error (Stratford & Goldsmith, 1997). Both male and female participants exhibited a similar range of *SEM*s for their handgrip strength measurements. In absolute terms, the highest *SEM* was of the dominant hands of $LONG_2$ males; a 2.8 kg f change would be needed from a mean grip strength of 44.1 kg f before the change could be considered to be beyond measurement error and therefore a real change in grip strength. However, the greatest *proportionate* change would be needed for the female $SQUARE_2$ group, who had a combination of lower mean strength and poor strength test reliability; a force change of 2.6 kg f would be required from a mean grip strength of 27.9 kg f before a real change in grip strength could be inferred for this group. This is a relative change of almost 10% in grip strength, as compared to the change of 6% in the male group. Because the *ICC* reliability values represent the proportion of the total variance in a set of scores that is associated with true differences between participants, about half of the difference in strength scores for this group of females represents error rather than true differences.

All the teenagers in the current study took part with good will and had nothing to gain from either a better or worse performance at retest, so explanations for the reliability values in terms of motivation differences can be excluded. One possible account of the significant finding is that, for females, only average or relatively longer-shaped hands are suited to the shape of the GripTrack™ and other Jamar-like handgrip dynamometer handles. Relatively squarer-handed females may be less able to conform to, or "map" onto the dynamometer handle, and may have experienced a greater amount of discomfort, compared with other females, during attempted maximum grip strength trials. Stephens, Pratt and Michlovitz (1996) found that 27% of their research participants regarded the Jamar™ dynamometer as uncomfortable to grip. If this was the case here, variability in the experience of discomfort may have been negatively correlated with reliability of grip strength effort. If it was the females with relatively squarer hands who experienced an uncomfortable grip posture more frequently than the others, their performance may have been less consistent. However, no student was allowed to continue with a grip strength test session if they reported experiencing pain that they perceived was reducing their grip strength performance that day. Alternatively, those females with relatively squarer hands may have had other unidentified physical, or psychological characteristics that led to their observed poorer reliability values.

These results have clinical significance when considering sincerity of effort issues and the design of grip strength dynamometer handles. If during hand therapy sessions a female teenage patient is giving inconsistent effort during repeated grip strength testing over a number of weeks, data from the current study suggest another possible explanation other than she is: (a) a malingerer, (b) affected by medications, or (c) has a psychosomatic disorder. If the individual concerned has a squarer hand relative to other females, this alone may disadvantage her in being able to give consistent performance over time.

IMPLICATIONS AND FUTURE STUDIES

Whether these results can validly be applied to adults with hand injuries requires investigation. Thus this hand-shape study could be extended to an injury-free adult population. Jamar^{TM} and Jamar^{TM}-like handgrip dynamometers, such as the computer-linked GripTrack^{TM} are commonly used in rehabilitation clinics throughout the world (Bohannon, 1991; De Smet & Vercammen, 2001; Kerr et al., 2006). Further research should be undertaken to investigate whether there are other handle shapes for isometric tools that do not disadvantage the female with relatively squarer hands. Comparisons could also be made with isokinetic tools. If Jamar^{TM}-like handle geometry is shown to be a factor related to the decreased reliability associated with these squarer-shaped hands then a design modification is needed for these existing dynamometers that permits equivalent reliability across different hand shapes.

CONCLUSION

Results suggest that hand-shape does not significantly influence the ability to generate maximal isometric grip strength, but that females with squarer-shaped hands are less able to reliably generate the same grip strength force readings using isometric handgrip dynamometers over retest time intervals of 4 weeks or less.

CHAPTER 11

CONCLUSIONS AND RECOMMENDATIONS

The testing of upper limb strength has been conducted for a variety of reasons for over three hundred years. The tools and methods used have developed from large uncalibrated, unstandardized analogue spring scales (Pearn, 1978a) to minute computerised polymer sensors (Kargov et al., 2004). It seems likely that in the 21st century isometric grip testing maximal grip strength will continue to be used as an indicator of change in hand function over time (Bohannon, 1998; Chimes, Foye & Braddom, 2006; Mannerkorpi et al., 2006), as it is apparent that this type of testing is considered valid by medical and legal authorities, and for this reason it will remain as an accepted measure.

On consideration of the literature on grip strength testing, personal experience as a practicing hand therapist and involvement with medico-legal proceedings for compensation, it was clear that what was needed to be done was to improve its reliability and hence its value as a measuring tool and predictor of the recovery of hand function. To do this it was necessary to identify the controlling variables that influence grip strength outcomes and to quantify these influences.

OUTCOMES OF THE RESEARCH

The outcomes presented in this thesis have identified some previously unknown variables affecting grip strength assessment in a group of teenagers, shown to be representative of Australian and international teenagers. The influences of the new and already known variables were quantified, prediction equations to assist clinicians in estimating grip strength in this population were formulated, and a set of normative anthropometric and grip strength data for this age group was established.

In achieving these major outcomes, other developments were also made.
A current literature review of the tools used to assess grip strength was conducted. A modified Edinburgh Handedness Inventory (Oldfield, 1971) was developed, trialed

with adults and teenagers, and found to be an improvement on the original in terms of reliability and content validity for the local population in this current era. The prevalence of past hand and forearm fractures in urban Australian teenagers was established. New height and weight norms for Brisbane-based teenagers were recorded. Factors that affected the generation of maximal isometric grip strength by these teenagers were identified. The most influential factors included age, gender, hand surface area, height and hand dominance.

Gender-split levels of reliability of grip strength measurements of the dominant and non-dominant hands were documented for retest over two different time periods. The level of reliability of the grip strength values did not alter significantly between a retest of one or four weeks. However the reliability values for the males in the various age and time interval groups ranged from $ICC = 0.94$ to 0.99 and were significantly higher than those for the females (*ICC*s from 0.72 to 0.82). When sorted into 3 handedness groups the reliability of the two hands was significantly greater in right dominant (dominant hand $ICC = 0.97$, non-dominant hand $ICC = 0.96$) and mixed-dominant teenagers (dom hand $ICC = 0.96$, non-dom hand $ICC = 0.96$) than in left-handed teenagers (dom hand $ICC = 0.47$, non-dom hand $ICC = 0.36$), although this result is tempered by the fact that the sample of left-handed teenagers was small (n=10).

The reliability of the grip strength ratio of the two hands in teenagers over a one to four week interval was determined for this sample. For 73% of the sample, the grip strength ratio did not change by more than 10% over either a one or four-week period. Age, gender and retest time interval did not influence this change. The types of statistics used to perform the analyses of the grip strength ratio were shown to be crucial to giving a complete picture of this motor behaviour. The minimal difference to be exceeded (*MDE*) before a change in grip strength could be considered to be significant varied depending on how long the retest interval was and the age and gender of the teenagers.

Hand shape of females was found to be the most striking factor in assessing the reliability of grip strength performance. If a female with squarer-than-average shaped hands was retested her repeat grip strength reliability value was often low, yielding *ICC*s of 0.48 for her non-dominant hand.

RECOMMENDATIONS FOR FUTURE RESEARCH

The modified Edinburgh Handedness Inventory (with the two new items) needs further testing with adults and teenagers to confirm its reliability and validity. Greater numbers of left-handed participants and square-handed females are required for repeat grip strength tests. Three repetitions of the grip strength tests with intervals of one to four weeks would be helpful to confirm reliability patterns. Past and current grip strength research strongly suggest that grip strength testing tools need to be re-designed to overcome the problems of hand shape bias and pain and discomfort that the Jamar™-like dynamometers appear to present to these teenagers, and then these newer models rigorously tested for reliability with defined populations.

CLINICAL APPLICATIONS

From the data collected for this project it is apparent that teenagers are not a “pristine” and injury-free group. Sixty-eight percent of 17-year-old males and 52% of 17-year-old females had sustained at least one upper limb fracture over their lifetime. Any permanent impairment from these fractures, or other upper limb injuries, need to be identified and taken into consideration when estimating loss of grip strength for rehabilitation and compensation purposes.

Data obtained here show that when young females with squarer-than-average shaped hands are grip strength tested they may not give consistent grip strength readings from one week to the next. It should therefore not be assumed that these females are malingering, or inconsistently feigning sub-maximal efforts. The poorer reliability of all the females in the retest situation could be due to a number of factors including levels of previous experience and the cortically mediated sense of effort in the production of a maximum voluntary isometric grip strength force. Discomfort and pain with use of the hand grip dynamometers has been a long standing problem (see history of grip strength tools in Chapter 2) and the lower *ICC*s of the females are evidence to confirm that it is a serious problem that could bias the interpretation of female grip strength performance.

Some important benchmarks are now available following this project. Australian clinicians can confidently use the grip strength normative databases presented here to compare their urban teenagers for grip strength and grip strength reliability parameters. If their male teenage patients have one hand injured, the grip strength of the other hand can be used to accurately estimate the pre-injury strength of that hand, with as much as 90% of the variance accounted for. With teenaged females only 70% of the variance in the grip strength of the dominant hand can be accounted for by the grip strength of the non-dominant hand. However, an increase to 74.3% of the variance in grip strength of their non-dominant hands can be accounted for by knowing the grip strength of their dominant hands, and adding the hand's surface area to the relevant prediction equation.

If clinicians have a teenage patient with bilateral upper limb injuries, they can use the given prediction models to estimate the pre-injury grip strength of both hands based on age, gender, height and hand surface area, with 34 to 63.5% of the variance between the actual and predicted grip strength accounted for by the input of the value of these parameters.