Journal of Scientific Exploration, Vol. 22, No. 4, pp. 495–515, 2008 0892-3310/08

Changes in Physical Strength During Nutritional Testing

CRAIG F. BUHLER

ChiroMAT Clinic, 447 North 300 West, #5, Kaysville, UT 84037

PAUL R. BURGESS

Department of Physiology, University of Utah School of Medicine,
420 Chipeta Way, Suite 1700,
Salt Lake City, UT 84108-6500
e-mail: p.r.burgess@utah.edu

EARL VANWAGONER

Quest Medical Group, Inc., 9304 Teton Estates Drive, West Jordan, UT 84088

Abstract—The ability of a patient to effectively stabilize (lock) and thus prevent rotation of his/her joint against forces applied by one of the authors (C.F.B.) appears to vary with the nature of the substance with which the patient is in proximity. C.F.B. believes the patient's nutritional needs can be evaluated by determining which of several substances is most effective in strengthening the weakened patient. The purpose of this study is to determine why such changes in physical strength occur. Bottles of pills were placed one after another on the supine patient's abdomen and chart recordings of force vs. time were made for the most effective strengthening agent and for an inactive placebo substance as the patient attempted to resist shoulder flexion. In the first set of trials, the tester was unblinded and the patient blinded. These same substances next were placed in opaque cloth bags so that both the tester and the patient were blinded and charted again. Finally, recordings of force vs. time again were made for the strengthening agent and the placebo substance, with the tester unblinded and the patient blinded. The tester, C.F.B., favored a signature hypothesis in which some electromagnetic emanation from the beneficial substance made the patient strong. Another author, P.R.B., favored a mental influence hypothesis, whereby C.F.B. made the patients strong through a subconscious mental influence reflecting his belief that a particular substance would be beneficial. When C.F.B. was unblinded, the patients were weak (yielded) in every placebo trial and were strong (locked the joint) in every trial with the active agent. When C.F.B. was blinded, patients again were weak with one compound and strong with the other; but C.F.B. correctly identified the active and placebo substances in only 8 of 27 patients. These findings support the mental influence hypothesis. Neither C.F.B. nor the patients were consciously aware of any mental interaction. When strong, individual patients could hold against peak forces on average 20–90% higher than the forces that caused them to yield when weak.

Keywords: motor function—nutritional testing—mental influence—electromagnetic signature

Introduction

Manual muscle testing has long been a part of patient evaluation. It is used by physiatrists, orthopedists, neurologists, chiropractors, osteopaths and physical therapists, among others. The patient is positioned so that the muscle or muscle group to be tested causes trunk or limb rotation in a particular direction. The tester then applies a counterforce as the patient makes a maximal effort to produce the desired rotation. Beasley (1956, 1961) recognized patient motivation increased when enough force was applied by the tester to cause the limb to rotate in the direction opposite that intended by the patient. This provides a defined endpoint for both the patient and tester.

Beasley stressed the desirability of making quantitative measurements of the forces produced by the patients. More recent studies (Burgess & Jones, 1997; Jones & Burgess, 1998) showing subjects can produce different amounts of force at a given effort level (effort-force rescaling) underscore the importance of making quantitative force measurements.

In the 1960s, manual muscle testing began to be used in unexpected ways (Diamond, 1979; Goodheart, 1966, 1976; Hawkins, 2002; Walther, 1976). For example, it was reported that simply holding a presumably harmful substance like lead in the hand made almost all subjects weak when any of a number of muscles were tested, whereas subjects were strong when holding a presumably beneficial substance like vitamin C. The tests most commonly were done using the shoulder joint, with the tester pressing down at the wrist on the outstretched arm of the standing patient, the starting position of the arm being parallel to the floor. The patient held the substance with the other hand and whether s/he knew what was being held made no difference to the outcome (Hawkins, 2002).

This led to the hypothesis that muscle testing could be used to assess the nutritional needs of patients, what substances they might be allergic to, the appropriateness of medications taken, etc. If it were true that patients would be weakened by holding substances harmful to them and strengthened by holding beneficial substances, the medical and scientific impact would be considerable.

Using quantitative force measurements, one of the authors (C.F.B.) has verified that the strength of a blinded patient can vary with the nature of the substance with which the patient is in proximity. The goal of this experiment is to determine why such changes in motor function occur.

C.F.B. favors the hypothesis that different substances have distinctive electromagnetic signatures that change the patient's strength. This proposition is called the signature hypothesis.

Another author (P.R.B.) thinks that C.F.B. changes the patient's strength with a subconscious mental influence that reflects C.F.B.'s beliefs, desires and expectations. This is called the mental influence hypothesis.

To test these hypotheses, blinded patients who were weak following 'diagnosis' of a medical problem were evaluated with different nutritional substances until one was identified that produced a strong response (the active or

strengthening substance). Another substance was identified that did not improve the strength of the already weak patient (the placebo substance).

The signature hypothesis predicts that when C.F.B. is blinded, the active substance will strengthen and the placebo substance will not. The mental influence hypothesis predicts one substance will strengthen and the other will not, but C.F.B. will not be able to reliably identify the active and placebo substances. The mental influence hypothesis makes this prediction because C.F.B. knows in advance he will be given both substances to test on each patient. In order for his signature hypothesis to be verified, he must find that one substance strengthens and the other does not. However, he will not know which is the active substance and which is the placebo.

Methods

Patient Selection

Of the 27 patients who participated in this study, only 2 were not currently being seen at the ChiroMAT Clinic. Any actual or potential patient was considered eligible. They were asked if they would like to be in an experiment designed to test how nutritional substances influenced motor function. In exchange, they would receive a nutritional evaluation and a nutritional supplement free of charge. All agreed and were scheduled to appear on 1 of 2 days (see Procedure).

The participants ranged in age from 17 to 79 years (mean 49.9). Fourteen were women and 13 were men. All but 2 were right-handed. One right-handed patient with a painful right shoulder was tested on the left. All other patients, including the 2 who were left-handed, were tested on the right. The Institutional Review Board of the University of Utah has determined this experiment to be exempt because it does not expose the patients to anything other than regular chiropractic care.

Measuring Strength and Stamina

Manual muscle testing was used in this study to evaluate the nutritional needs of patients. One of the authors (C.F.B.) did all of the evaluations. The patient was supine with the pronated right arm held straight upward at a right angle to the frontal plane. Force was applied to the patient's wrist by the examiner with a force transducer (Hoggan FET System) interposed between the examiner's hand and the patient's wrist. The force was applied at right angles to the same place on the back of the wrist (± 1 cm), thus tending to flex the shoulder with constant leverage. The patient was instructed to keep the back of the shoulder flat on the table and the elbow straight and locked while making a maximal effort to resist shoulder flexion. Any attempt by the patient to shift the body out of the test position during the application of force should be seen as an inappropriate response and the test should be discarded. This is a critical component of any muscle test as it indicates recruitment and leads to increased

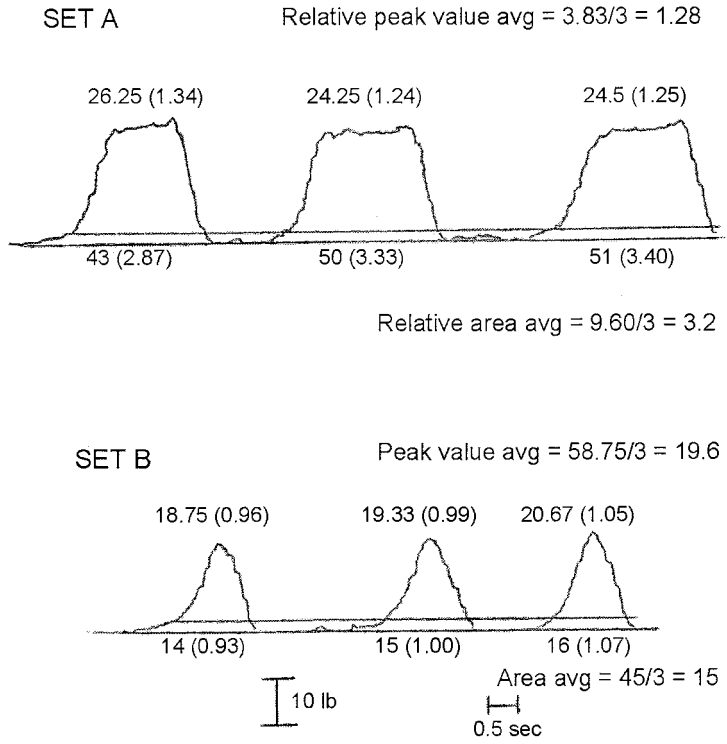

Fig. 1. This is a record of force vs. time obtained with the tester unblinded and the patient blinded. The active agent was on the patient's abdomen during set A and the placebo substance during set B. When the placebo was present (set B), the patient yielded in every trial, the shoulder flexed about 10° and the force declined. When the active substance was present, the patient was able to lock the joint, even though the peak force was higher, and did not yield in any trial. The numbers above the tracings are peak force values in pounds, with those in parentheses indicating the values relative to the mean of set B. The numbers below the tracings are for the areas enclosed by the force vs. time tracings. One unit = 0.75 lb-sec. The force-time areas were measured from a baseline (upper horizontal line) that was equal to 10% of the mean peak force value in set A. The numbers in parentheses indicate the values relative to the mean area in set B. When the active substance was present, the mean peak force was higher by a factor of 1.28 and the mean force-time area was greater by a factor of 3.20.

adaptive strength and faulty results. The patient was alerted in advance of each trial that force was about to be applied.

Muscular strength was measured after a bottle of pills of the nutritional substance to be tested was placed on the patient's abdomen. The bottle, as it was moved to and rested on the abdomen, was out of the patient's line of sight; the back of the patient's head and shoulders always was against the examining table during muscle tests. Since C.F.B. did not tell the patient what supplement he was testing, the patient was blind.

Figure 1 is a record of force vs. time obtained with the tester unblinded and the patient blinded. The active agent was on the patient's abdomen during set A

and the placebo substance during set B. The way in which the patient's forces are measured is designed to emphasize the improvements in strength that can occur in the clinic. When the patient is weak, sufficient force is produced by the examiner to flex the patient's shoulder 10–15°; i.e., the patient yields in every trial (set B). When the patient is strong, the force is increased to a higher level, but not so high that the patient yields, and is sustained for about 2 sec, at which time the examiner releases the force (set A). This has the advantage that it is less likely to injure the patient, but has the disadvantage that it underestimates the patient's ability to produce force because the tester terminates the trial before the patient yields.

In order to evaluate these force measurements, it is helpful to understand something of the biomechanics of strength testing and the design of the force transducer. The force acting back on the force transducer from the patient's wrist equals the force with which the examiner is pressing against the wrist. Gravitational effects on the transducer are small and can be ignored. The only forces relevant for flexion of the arm at the shoulder are those that act at right angles to the wrist in the direction producing shoulder flexion torque. The force transducer accurately measures this force by having a side-load rejection ratio of more than 100 to 1 and by being insensitive to the point on the contact surface where the force acts.

The next question is what features of the patient's force response should be measured? Two have been chosen: the peak force and the force-time area. The peak force is a widely used measure of physical strength and has been shown to be reliably evaluated with hand-held dynamometers (Bohannon, 1995, 1997a; Brinkmann, 1994; Hsieh & Phillips, 1990; Mulroy et al., 1997; see Sapega, 1990, for a general discussion of force measurement in the clinic and Bohannon, 1993, for a list of the 132 studies in which hand-held dynamometers had been used up to 1993). The precision of hand-held dynamometry compares favorably with fixed dynamometers (Brinkmann, 1994; Magnusson et al., 1990; Stratford & Balsor, 1994) and hand-held dynamometers display good consistency when a patient is repeatedly tested (Bohannon, 1986, 1997a,b; Hosking et al., 1976; Hsieh & Phillips, 1990) so long as there is a well-defined endpoint for the test and the tester can easily break the stabilized patient (Agre et al., 1987; Byl et al., 1988; Mulroy et al., 1997).

The force-time area is the area under the force-time curve and is influenced by both the amplitude of the force and its duration. It has been used in the past as a global measure of motor output (Nicholas et al., 1978; Ryan & Agnew, 1917). The force-time areas were measured with a planimeter.

Both peak force and force-time area measures have a straightforward interpretation when the patient resists displacement (Figure 1, set A). When the patient yields and the arm moves (Figure 1, set B), some of the force measured by the transducer is that required to produce angular acceleration of the limb. What is important in the present context is that these inertial forces add to the force measured in the placebo trials and so work against the hypothesis that

therapy strengthens the patient. It is especially important to minimize Type I error when doing research that is potentially controversial and no attempt has been made to correct either peak force or force-time area measurements for inertial forces.

Another factor increasing the measured force during the placebo trials is the well-known increase in muscular force that occurs when a contracting muscle is stretched even though there has been no increase in the neural drive to the muscle. Animal studies have shown that these increases can be as great as 10–20% of the prestretch isometric force (Joyce et al., 1969). Once again these forces increase the placebo break values and so work against the hypothesis that therapy strengthens. No attempt was made to correct either peak force or force-time area measurements for these eccentric enhancements in muscular force.

Stretch reflexes also add to the muscular forces produced during the joint rotation that occurs during the placebo trials. These forces would be reduced for a given fusimotor drive during the trials in which the patient resists displacement. Stretch reflex force also works against the hypothesis that therapy strengthens and no attempt has been made to estimate or correct either peak force or force-time area measurements for these forces.

In summary, it can be argued that peak force and force-time area measurements of the placebo trials (Figure 1, set B) should be restricted to the portion of the trial before the limb begins to move. This would ensure that the placebo measurements are comparable to the measurements of the trials with the active agent, which cover only times when the patient has control of the load and is successfully resisting displacement. Another argument for this approach is that the patients were instructed not to let their arm be displaced and to make a maximal effort to prevent this. How much they should resist after they lose control of the load was not specified. Although patients were not encouraged to resist after movement began, their behavior will have varied with their interpretation of the task and their motivation. As a consequence, whether the force peaks as movement begins or later may vary from patient to patient. In any case, once the force begins to fall in a continuous fashion, it is safe to assume that movement is underway. Hence, measurements of force-time area in the placebo trials should be restricted to the time between when the force exceeds threshold and when it peaks.

The counter-argument to these ideas is that to minimize Type I error, the methods used should always work against the hypothesis that the patient is strengthened by therapy. For this reason, we have included the entire placebo force response, including the declining phase of the force, in the force-time area. (Note that a threshold is set below which the force is not measured in order to exclude baseline fluctuations; see Figure 1.) The effect of including the declining phase of the force is to increase the placebo force-time areas and thus reduce the magnitude of any increase of the active-agent trials over the placebo trials. This increase in the placebo trials is somewhat diminished because some

of the patient's muscular force has to be diverted to support the arm against gravity as the shoulder flexes and so is not registered by the force transducer. No correction has been made for this gravitational effect.

A 2-tailed unpaired t test (one that does not assume equal variance) was used to evaluate the statistical significance of changes in the magnitude of peak forces and force-time areas. Lines were fit to data points using linear regression analysis.

Procedure

The experiment was done on 2 different days; 10 patients were tested on the first day and 17 on the second.

Patients arrived at the clinic at an appointed time. They read a form describing the possible risks and benefits of participating in the experiment, and all indicated their willingness to participate by signing the form.

The patient was escorted to an examining room where C.F.B. and P.R.B. waited. The patient's motor status was evaluated by C.F.B. and all patients tested strong on arrival. C.F.B. then searched for some indication of an incipient or actual health problem using methods honed over many years of practice. It is beyond the scope of this paper to describe these methods in detail. Suffice it to say that from the perspective of Western medicine, they would not be expected to change physical strength.

Within 2–10 min, every patient was diagnosed with a medical issue that was revealed when s/he tested weak. Typically, 4–6 bottles, each containing a different nutritional supplement appropriate for the diagnosed medical condition, were placed one by one on the patient's abdomen until the most effective strengthening agent was found. A bottle containing a placebo substance also was identified, which, when placed on the patient's abdomen, did not improve the patient's ability to resist shoulder flexion; i.e., the patient remained weak. The patient's strong response with the strengthening bottle on the abdomen was then charted. The patient's weak response was next charted with the placebo bottle on the abdomen. Three trials were done with each substance. These tests constitute the first part of the nutritional experiment. The patient was blinded and the tester was unblinded.

In the second part of this study, the tester (C.F.B.), P.R.B. and the patient were blinded. The proper dosage of the strengthening agent was determined first by opening the pill bottle, placing the inverted bottle cap on the patient's abdomen, adding 1 pill to the bottle cap and testing shoulder flexion to see if the patient was as strong as with the entire bottle of pills. If necessary, another pill was added and the test repeated. The maximal dose for this group of patients was 6 pills; the minimum was 2.

P.R.B. took a glass bottle holding the appropriate dose to another room and waited outside the closed door while another individual (the blinder) placed enough placebo pills into an identical glass bottle so that the 2 bottles had

a similar weight. The bottles (9.5 mm tall, 3.5 mm diameter, 68 gm weight capped and empty) were capped and placed into 1 of 2 cloth bags (15 mm wide, 10 mm high, with a zipper along the top, 17.7 gm weight). One bag was black and the other, tan. Whether the active agent would be in the black or the tan bag was predetermined using a random number table. Randomization was violated near the end of the assignment so that the active agent was distributed between the black and tan bags as equally as possible.

The 2 bags were weighed (Sartorius PT120 portable electronic balance). The amount of placebo was adjusted until the bags weighed almost the same. The largest difference in bag weight was 0.7%; the mean difference was 0.1%.

A different bottle was used to hold the strengthening substance for each patient. The same bottle was used for the placebo substance for all but 1 patient. Only 1 patient required a different placebo and a different bottle was used for that patient.

When the blinding was complete, the blinder opened the door to her room while holding both bags in the other hand and gave them to P.R.B. He returned to the examining room with the bags. The active agent was in the black bag for 19 patients and in the tan bag for 18 patients. C.F.B. always found one bag to strengthen the patient and the other not to strengthen. For 15 patients, he found the tan bag to strengthen. The remaining 12 patients were strengthened by the black bag. These outcomes were charted with 3 trials for each bag. These tests with everyone in the examining room blinded constitute the second part of the nutritional experiment.

In the third part of the experiment, the strengthening and placebo bottles used in the first part were retested with C.F.B. unblinded and charted with 3 trials for each.

The patient was given the bottle of strengthening pills minus those in the cloth bag. The cloth bags were returned to the blinder by P.R.B. Then P.R.B. brought the next patient to the examining room, etc. The average duration of a session was 14 min; the shortest was 8 min and the longest was 22 min.

C.F.B. was not told whether he had accurately identified the active and placebo substances until all patients had been tested. The person analyzing the data did not know whether C.F.B. had been correct or incorrect in the blinded part of the experiment.

Summary of Procedure

1. Patient's motor status is evaluated. All patients test strong.
2. Incipient or real medical problem is revealed when patient tests weak.
3. Appropriate nutritional substances are sequentially placed on patient's abdomen until one is found that makes the patient strong.
4. Placebo substance is identified that leaves patient weak.
5. Chart recordings of force vs. time are made for strengthening and placebo substances with C.F.B. unblinded and the patient blinded.

6. Step 5 is repeated with both C.F.B. and the patient blinded.
7. Step 5 is repeated with C.F.B. again unblinded and the patient blinded.

Normalization of the Data

The experiment consisted of 3 parts for each patient. In each part, the patient was tested with a strengthening substance and a placebo substance, and 3 records of force vs. time were made for each substance (Figure 1). In the first and third parts (tester unblinded, patients blinded), the patients always were able to resist flexion of the shoulder joint when the strengthening substance was present and always yielded in the presence of the placebo. In the second part (both tester and patients blinded), the patients also were strong and weak, but C.F.B. correctly identified the active agent in only 8 of the 27 patients.

There were 6 sets of records for each patient. Each set consisted of 3 trials: a weak and a strong set for the first part, a weak and a strong set for the second part, and a weak and a strong set for the third part. The 3 parts were normalized separately. Using patient 1 to illustrate the method, the individual peak-force values in the weak and strong trials of the first part were divided by the mean value of the peak forces when the patient was weak (Figure 1). This allows each peak force in the weak and strong trials of the first part to be expressed relative to the mean value of the weak trials. The relative mean for the weak trials, therefore, will equal 1; and the relative mean for the strong trials will be larger than 1 by a factor that indicates the magnitude of the strengthening effect. This puts all patients on the same footing regardless of strength. The force-time areas of patient 1 in the first part were normalized in the same way (Figure 1). This procedure was repeated on the data of patient 1 for the second and third parts. This allows the relative changes in peak forces and force-time areas for patient 1 in all 3 parts of the experiment to be compared. The data for the other 26 patients were normalized in the same way.

Results

Comparison of Unblinded and Blinded Data

Figure 2 shows the data from the first part of the experiment (tester unblinded, patients blinded). In this figure, the force-time area for each trial has been plotted as a function of the peak force for the same trial. Open circles indicate placebo tests where the patient yielded in every trial. Filled circles are for the strengthening substance; the patient held in every trial. That the patient did not yield increases the force-time area for a given peak force and magnifies the effect of increases in peak force on the force-time area, increasing the slope of the relationship between force-time area and peak force. In the presence of the strengthening substance, the mean peak force increased by a factor of 1.49 ($p = 9.22 \times 10^{-35}$, unpaired t test, 2-tailed) and the mean force-time area by a factor of 4.28 ($p = 1.14 \times 10^{-37}$).

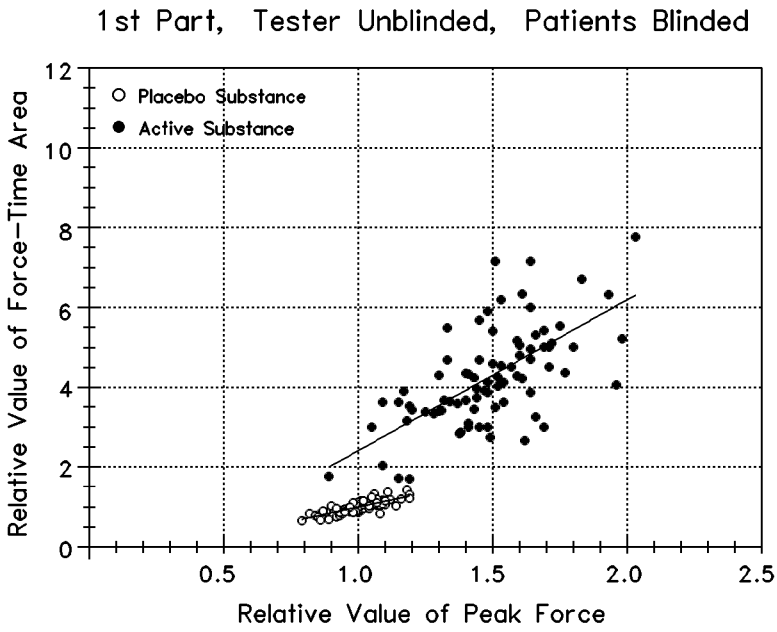

Fig. 2. Force-time area is plotted against peak force for each trial in the first part of the experiment. The tester was unblinded and the patients were blinded. Each of the 27 patients is represented by 3 trials for the placebo substance and 3 trials for the active substance. The open circles show the placebo tests and the filled circles the tests with the active substance. The patients yielded in every placebo trial and held in every active trial. The mean peak force increased when the active agents were present by a factor of 1.49 ± 0.23 ($p = 9.22 \times 10^{-35}$, unpaired t test, 2-tailed) and the mean force-time area increased by a factor of 4.28 ± 1.29 ($p = 1.14 \times 10^{-37}$). The mean peak force and force-time area values for the placebo tests are both 1.00 because the data were normalized. The standard deviations are ±0.087 and ±0.16 respectively. Linear regression analysis gave a slope of 1.45 for the placebo data with a correlation coefficient of 0.81, and a slope of 3.78 for the active substances with a correlation coefficient of 0.66.

Figure 3 shows data from the third part of the experiment (tester unblinded, patients blinded) similarly plotted and using the same bottles of strengthening and placebo substances. In the third part, the mean peak force increased by a factor of 1.50 in the presence of the strengthening agent ($p = 5.75 \times 10^{-28}$) and the mean force-time area by a factor of 3.97 ($p = 1.07 \times 10^{-39}$). The increases in mean peak force are almost identical in the first and third parts of the experiment; the increases in mean force-time area are larger in the first part, but not significantly so ($p = 0.10$). The increases in force-time area are more variable in the first part than in the third part (SDs of 1.29 and 1.07 respectively), but the difference is not statistically significant ($p = 0.10$, F test).

Figure 4A shows the data in the second part of the experiment (both tester and patients blinded) of the 8 patients for whom C.F.B. correctly identified the strengthening and placebo substances. The mean peak force increased by

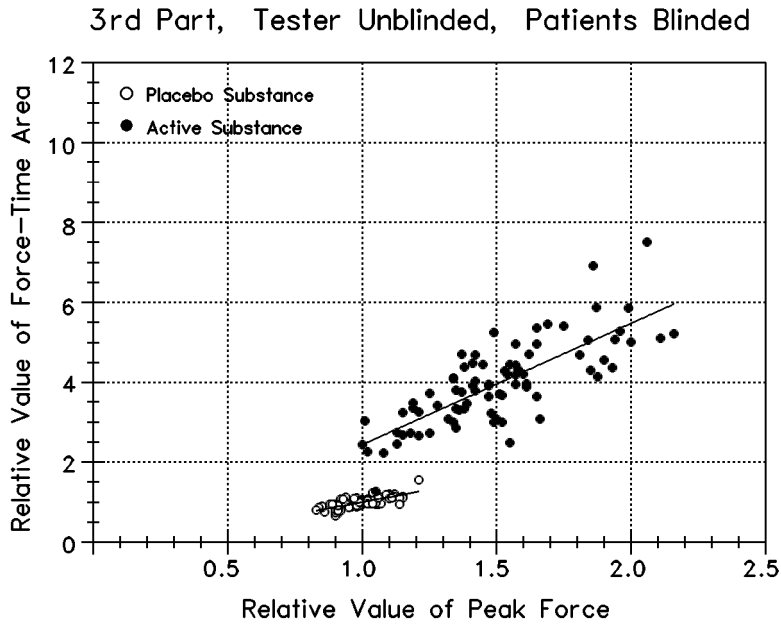

Fig. 3. The data from the third part of the experiment are represented as in Figure 2. The tester was unblinded and the patients were blinded. Each patient is represented by 6 trials, 3 with the placebo substance (the patient yielded in each of these trials) and 3 with the active substance (the patient held in each of these trials). In the presence of the active substance, the mean peak force increased by a factor of 1.50 ± 0.27 ($p = 5.75 \times 10^{-28}$) and the mean force-time area increased by a factor of 3.97 ± 1.07 ($p = 1.07 \times 10^{-39}$). The mean values for the placebo tests are 1.00 because the data were normalized. Standard deviations are 0.082 for peak force and 0.14 for force-time area. The slope of the placebo line is 1.25, with a correlation coefficient of 0.74. The slope of the line when the patients were strengthened is 3.04, with a correlation coefficient of 0.78.

a factor of 1.58 in the presence of the strengthening agent ($p = 2.02 \times 10^{-11}$) and mean force-time area increased by a factor of 4.37 ($p = 2.54 \times 10^{-12}$).

Figure 4B shows the 19 patients for whom C.F.B. incorrectly identified the strengthening and placebo substances. Here the mean peak force increased by a factor of 1.52 in the presence of the placebo substance ($p = 3.61 \times 10^{-19}$) and the mean force-time area increased by a factor of 4.16 ($p = 2.51 \times 10^{-22}$). The peak-force and force-time-area increments were not significantly different depending on whether the substances were correctly or incorrectly identified ($p = 0.35$ and 0.53 respectively).

In Figure 5, the data for the correct and incorrect judgments in the second part of the experiment have been combined for comparison with the first and third parts of the experiment. Mean peak forces increased by factors of 1.49, 1.53 and 1.50 in the first, second and third parts respectively. The mean force-time areas increased by factors of 4.28, 4.22 and 3.97. None of these differences were statistically significant.

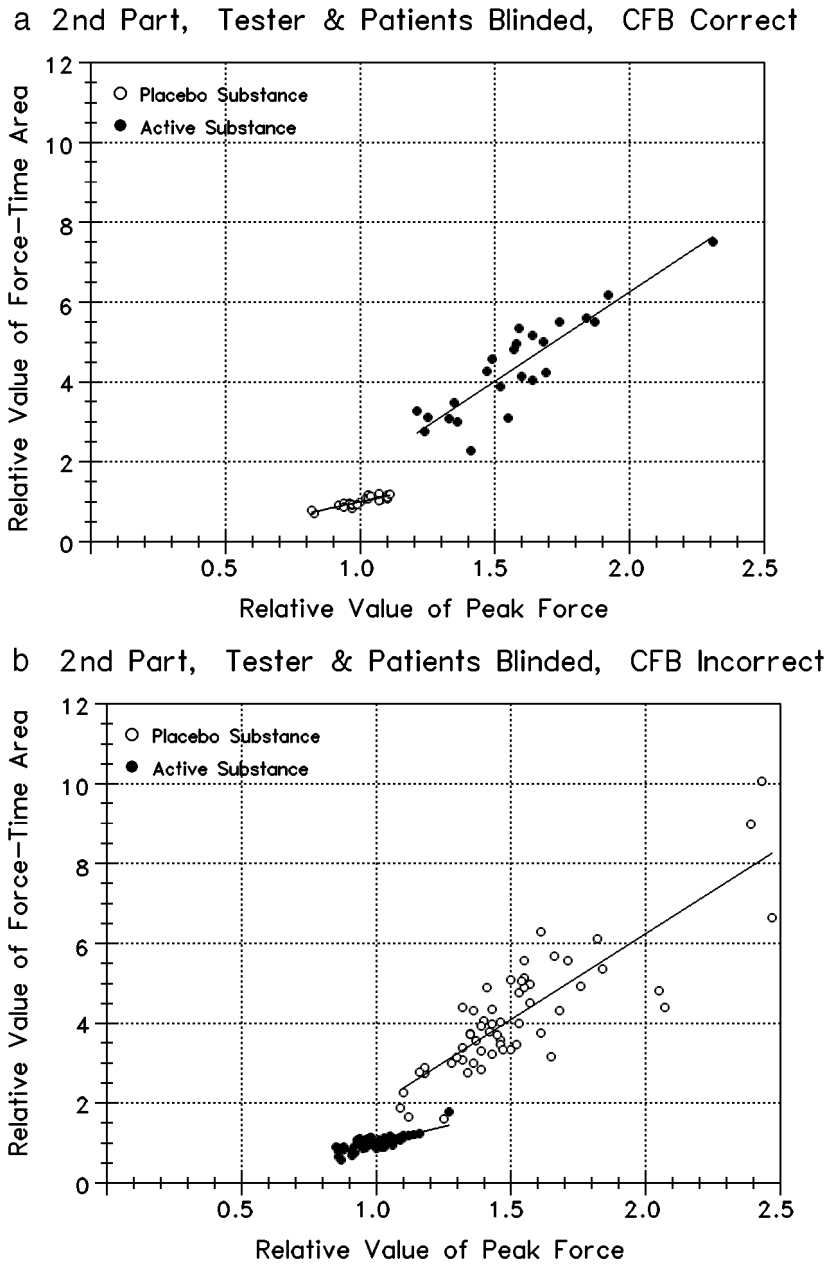

Fig. 4A. These are the 8 patients for whom C.F.B. correctly identified the placebo and active substances in the second part of the experiment (both tester and patients blinded). Again, there were 3 trials per condition with the patient yielding in every placebo trial and holding in every trial with the active agent. The mean peak force increased by a factor of 1.58 ± 0.25 when the active

The data for all 3 parts of the experiment have been combined in Figure 6A. The open circles indicate weak responses (trials where the patient yielded), which were always associated with the placebo when the tester was unblinded (first and third parts of the experiment); this also was true for 8 of the patients when the tester was blinded (second part of the experiment). The other 19 patients in the second part were weak in the presence of the active substance. The filled circles indicate strong responses (trials where the patient resisted joint rotation), which always were associated with strengthening substances when the tester was unblinded and in 8 of the patients when the tester was blinded. The other 19 patients were strengthened in the presence of the placebo when the tester was blinded. The patients when strong had peak forces 51% greater on average than when they were weak ($p = 8.92 \times 10^{-88}$) and the average force-time area was greater by 316% ($p = 3.12 \times 10^{-106}$).

Figure 6B shows the data in Figure 6A after averaging the weak and strong trials for each individual subject. Nine trials were averaged to obtain each of these mean values. The mean peak force and the mean force-time area equal 1 for each subject when weak because the data were normalized (large open circle). The filled circles indicate the mean values of the strong trials for each subject. The overall strengthening effect is robust when viewed patient by patient; individual patients, when strong, can stabilize peak forces on average 20–90% higher than the forces that caused them to yield when weak.

Discussion

Several hypotheses may be pertinent to our experiments.

Biomechanical Hypothesis

This hypothesis proposes that the strength of the patients does not change, but that patients when seemingly weak are caught off guard or somehow overwhelmed by the tester and that patients are treated with restraint when

---

←

agent was present ($p = 2.02 \times 10^{-11}$) and the mean force-time area increased by a factor of 4.37 ± 1.25 ($p = 2.54 \times 10^{-12}$). In the placebo trials, the means of the peak forces and force-time areas equal 1.00 because of normalization; the standard deviation is 0.079 for peak force and 0.13 for force-time area. The slope of the placebo trials is 1.46 (correlation coefficient 0.89) and the slope of the active-agent trials is 4.47 (correlation coefficient 0.89).

Fig. 4B. These are the 19 patients for whom C.F.B. incorrectly identified the placebo and active substances in the second part of the experiment (both tester and patients blinded). Here the patient yielded in all 3 trials when the active substance was present and held in all 3 placebo trials. The mean peak force was greater by a factor of 1.52 ± 0.30 when the placebo substance was present ($p = 3.61 \times 10^{-19}$) and mean force-time area was greater by a factor of 4.16 ± 1.52 ($p = 2.51 \times 10^{-22}$). The means of the peak forces and force-time areas in the active-agent trials equal 1.00 because the data were normalized; standard deviations are 0.084 and 0.17 respectively. When the patients were weak, the slope of the line is 1.65 (correlation coefficient 0.79). The slope for the patients when strong is 4.30 (correlation coefficient 0.84).

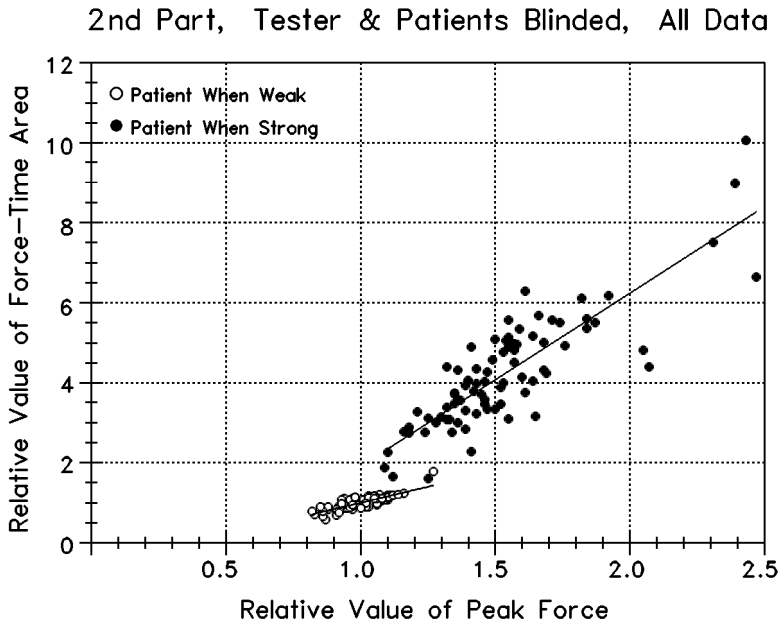

Fig. 5. The data in Figures 4A and B have been combined as there were no statistically significant differences between their peak forces and force-time areas. Patients when weak yielded in every trial and patients when strong held in every trial. The mean peak force is greater by a factor of 1.53 ± 0.28 ($p = 4.59 \times 10^{-29}$) in the strong trials and the mean force-time area is greater by a factor of 4.22 ± 1.44 ($p = 4.22 \times 10^{-33}$). The means for the weak trials are 1.00 with standard deviations of 0.082 and 0.16 for peak force and force-time area respectively. The slope of the line for the weak trials is 1.60 (correlation coefficient 0.81); and for the strong trials, the slope is 4.33 (correlation coefficient 0.85).

seemingly strong. Three things are important for this hypothesis. Was force applied to the patients when 'weak' with greater leverage or with the shoulder in a less favorable position? Did the trials of the patients when weak begin before they were ready? Were the patients when weak subjected to forces that built up more rapidly or to higher levels?

Great care was taken to apply the force at right angles to the same place on the patient's wrist in every trial. This was not difficult because of the obvious anatomical landmarks at the wrist. Accurate placement within ±1.0 cm can be achieved and produces an error of ±2.2% in muscular force estimation for a patient with an arm of 46 cm from shoulder to wrist. Almost all patients had arms longer than 46 cm. The patients were watched carefully to be sure that the shoulder was flat against the examining table and remained there during the tests. The patient's elbow was straight and locked in the extended position. Patients were not charted unless these conditions were met. The patients were always alerted by being told to resist just before the force was applied. Every chart recording used in this report was examined for evidence that the rate of rise

of the force was higher when the patients could not resist displacement. Figure 1 is typical; the tendency is for the rate of rise of the force to be greater when the patients were able to resist displacement. Figure 6B illustrates that individual patients when strong could hold against peak forces on average 20–90% greater than caused them to yield when weak.

Social Interaction Hypothesis

There was no observable change in C.F.B.'s manner that might explain whether a patient was weak or strong. P.R.B. accompanied C.F.B. whenever he was with patients and C.F.B. was uniformly agreeable and professional in his conduct. At no time were the patients when weak subjected to intimidation or distraction, nor were the patients when strong selectively encouraged.

Mechanical Interaction Hypothesis

According to this hypothesis, C.F.B. applies force in subtly different ways to the patient depending on whether he wants the patient to yield or resist. C.F.B. trains the patient to recognize these subtle mechanical cues so the patient will know what C.F.B. expects. Neither C.F.B. nor the patient is consciously aware of the existence of these mechanical cues or of the training process.

This hypothesis faces 3 major difficulties. (1) Naïve patients may test weak, for example, and a few seconds later test strong after an intervention thought to be beneficial. There is no time for training in such cases. (2) Even if the patients could be trained to yield or resist based on subtle mechanical cues, there is no reason for them to yield at lower forces than they can resist without yielding. Patients make a maximal effort in every trial, and this is particularly noticeable when one is forced to yield. As discussed in Methods, for a given motor drive patients would be expected to yield at higher forces than they can successfully resist. Thus this hypothesis fails to account for the observation that patients can be weakened and strengthened by C.F.B. (3) A careful search of the patient's force records yielded no evidence of a uniformly different way in which C.F.B. applied force when the patient was weak vs. strong. There is a tendency for the force to rise more rapidly when the patient is strong (see Biomechanical Hypothesis and Figure 1), but there are cases where the forces rise at similar rates and a few where the force rises more rapidly when the patient yields.

Placebo Hypothesis

This hypothesis attributes strengthening to a placebo response by the patient. It does not apply to these experiments because the patients did not have a placebo response; they always were weak when the placebo substance was present in the first and third parts of the experiment (tester unblinded, patients blinded). It is difficult for patients to develop the expectations needed for a placebo response when they cannot predict whether they will be weak or

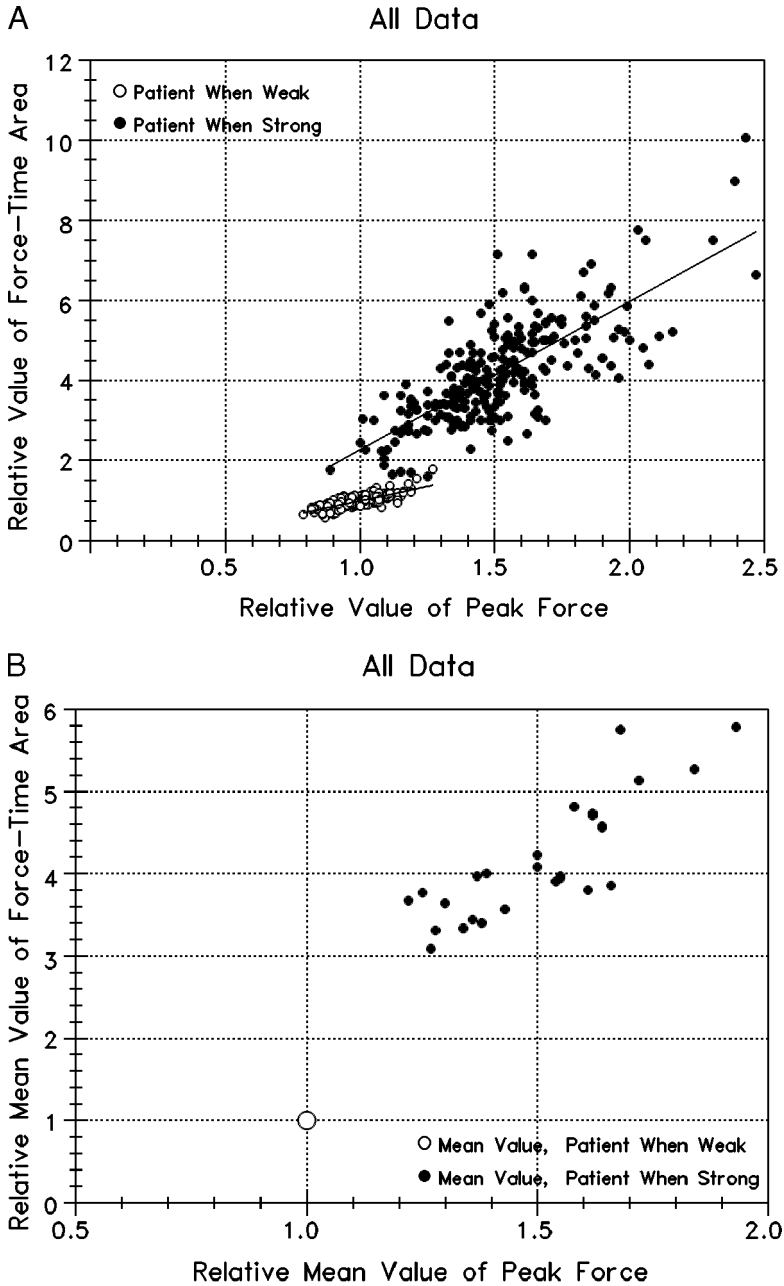

Fig. 6A. The data from all 3 parts of the experiment (Figures 2, 3 and 5) have been combined, since there were no statistically significant differences between their peak forces and force-time areas. Patients when weak yielded and patients when strong held in every trial. The mean peak

strong while being tested. From the point of view of the patients, the tester controls their physical strength.

Signature Hypothesis

This hypothesis proposes that a substance held in the hand or resting on the abdomen improves a patient's ability to resist joint rotation because the substance has an electromagnetic emanation that is strengthening. This hypothesis predicts C.F.B. when blinded will be able to accurately identify substances that strengthened the patient or had no influence on the already weak patient.

Mental Influence Hypothesis

This is a broad category of interaction that includes both conscious expressions of the will and subconscious mental manifestations of beliefs, desires, attitudes and expectations. This hypothesis predicts C.F.B. will find patients to be weak with one of the blinded compounds and strong with the other; this is the only outcome allowed by the signature hypothesis, which he favors. The mental influence hypothesis also predicts C.F.B. will not be able to reliably identify the blinded substances.

The active agent always strengthened the patients when the tester was unblinded, but the active agent strengthened only 30% of the patients when the tester was blinded. In 70% of the trials, the placebo strengthened the patients.

This result is clearly supportive of the mental influence hypothesis. The chi-square test can be used to determine the likelihood this result would have occurred by chance alone. Once the outcome with the active agent is specified, the outcome with the placebo also is known; therefore only the results with the active agent were tested.

The sample consisted of 27 patients; each was tested while blinded in both parts 1 and 3 of the experiment with the tester unblinded. These patients were strong in 54 out of 54 tests when the active agent was present. When the tester also was blinded, 8 patients were strong and 19 were weak with the active agent.

←

force is greater by a factor of 1.51 ± 0.26 ($p = 8.92 \times 10^{-88}$) in the strong trials and the mean force-time area is greater by a factor of 4.16 ± 1.28 ($p = 3.12 \times 10^{-106}$). The means for the weak trials are 1.00 with standard deviations of 0.083 and 0.15 respectively. The slope of the line for the weak trials is 1.44 (correlation coefficient 0.71); and for the strong trials, the slope is 3.79 (correlation coefficient 0.76).

Fig. 6B. The 9 weak trials have been averaged across the 3 parts of the experiment for each patient. The mean values are 1.00 for both peak force and force-time area because of normalization and are indicated by the large open circle. The 9 strong trials for each patient have been averaged also and are indicated by the filled circles.

These data were entered into a 2×2 contingency table. The chi-square is 45.8, $p < 0.0001$, and the mental influence hypothesis is supported.

The accuracy of the blinded tester differs from what would be expected if he were as likely to select the active agent as the placebo ($p = 0.034$, binomial test, 2-tailed). It is not known why in our study the blinded tester correctly identified the active agent only 30% of the time.

There was no awareness on the part of C.F.B. or the patients that any mental interaction had occurred either when C.F.B. was unblinded or blinded. It is for this reason that the mental influence is considered to be subconscious. For C.F.B. consciously to will a particular patient to be strong or weak would have been inappropriate because he believed that he simply was evaluating the motor status of the patient.

It is not known how one person's mind can act on another person's motor function. The mental influence hypothesis is consistent with an earlier study in which the ability of subjects to hold against the muscular forces of an experimenter apparently was influenced by the intentions of the latter (Burgess & Wei, 1984). It is only a modest extension of the interpersonal mental influence hypothesis to an intrapersonal version, which states that one's own mind can influence one's own motor behavior. We have been unable to find any paper in the peer-reviewed scientific literature that provides evidence for a direct mental influence of one person on the physical strength of another.

Other Related Studies of Motor Function

There have been a number of studies in which the more controversial aspects of muscle testing have been examined. They can be divided into those with force measurements and those without. Without force measurements, there is no way to rule out the biomechanical hypothesis discussed above. If the tester is simply overwhelming the patients/subjects when they are found to be weak, and is unaware of this due to effort-force rescaling (Burgess & Jones, 1997; Jones & Burgess, 1998; see also Hyman, 1999), muscle testing outcomes that at first seem surprising are relatively easily explained.

Studies in which force is measured and subjects are found to be strong and weak for no 'known' reason pose a challenge to orthodox science, even when the experiments are unblinded (Caruso & Leisman, 2000; Chorbajian et al., 1988; Diamond, 1979; Monti et al., 1999; Omura, 1979; Scopp, 1978). The explanation that has been favored so far is that the subject/patient has been changed in some way by the substance with which s/he is in contact, the lie just told, the reflex (acupuncture) point pressed on, etc., and that these patterns of strength and weakness can be used for medical diagnosis and lifestyle recommendations (Diamond, 1979; Goodheart, 1966, 1976; Hawkins, 2002; Walther, 1976). Another possibility, the one supported by our findings, is that the beliefs, desires, attitudes and expectations of the practitioner determine whether the patient/subject is weak or strong.

Which of these 2 alternatives is correct can be established by a research design in which a series of tests is done to establish that blinded patients are reliably weak under one set of conditions and reliably strong under another set of conditions. Once these nonintervention controls are completed, the tester is blinded and the tests are repeated.

Several carefully done experiments have not studied the ability of subjects to resist joint displacement at the hands of a tester, but have used the subject's ability to produce handgrip or elbow-flexion force as an indicator of motor performance (Arnett et al., 1999; Braud, 1989; Keating et al., 2004; Kendler & Keating, 2003; Radin, 1984). In this design, the tester has been replaced by an experimenter who presents various substances to the subject who then makes a maximal effort against a dynamometer. Unfortunately, no unblinded experiments were reported showing which substances were reliably associated with strong and weak responses. After blinding both the experimenter and the subject, substances thought likely to weaken and strengthen the subject had similar effects. These results are consistent with the mental influence hypothesis, but this hypothesis would only be relevant if consistent weak and strong responses were obtained for particular substances when the experimenter was unblinded.

The mental influence hypothesis explains how perceptually dramatic changes in one's ability to resist joint rotation can occur that may or may not be correlated with particular stimuli, depending on whether the tester is blinded. However, our experiment was done at 1 clinic by 1 tester with a small group of patients. More studies need to be done before the generality of the mental influence hypothesis can be accepted. Perot et al. (1991) gave subjects a form of muscle manipulation aimed at unloading muscle spindle receptors within the muscle and thereby creating weakness. The effect was present after blinding the tester. This may mean that the mental influence of the manipulator, who remained in the room, was influencing the subjects. Alternatively, this procedure may actually produce muscle weakness by spindle unloading.

Acknowledgments

We very much would like to thank John Stockton for his gift in support of this research. We also wish to thank the following colleagues for their thoughtful comments on this and earlier versions of the paper: Philip Bolton, Jeffrey Denburg, Max Haid, Ken Horch, Stephen Perle, Reed Phillips, Joel Pickar, Victor Wilson and Walt Woodbury.

References

Agre, J. C., Magness, J. L., Hull, S. Z., Wright, K. C., Baxter, T. L., Patterson, R., & Stradel, L. (1987). Strength testing with a portable dynamometer: Reliability for upper and lower extremities. *Archives of Physical Medicine and Rehabilitation, 68,* 454–458.

Arnett, M. G., Friedenberg, J., & Kendler, B. S. (1999). Double-blind study of possible proximity effect of sucrose on skeletal muscle strength. *Perceptual and Motor Skills, 89,* 966–968.

Beasley, W. C. (1956). Influence of method on estimates of normal knee extensor force among normal and postpolio children. *Physical Therapy Review, 36,* 21–41.

Beasley, W. C. (1961). Quantitative muscle testing: Principles and applications to research and clinical services. *Archives of Physical Medicine and Rehabilitation, 42,* 398–425.
Bohannon, R. W. (1986). Test-retest reliability of hand-held dynamometry during a single session of strength assessment. *Physical Therapy, 66,* 206–209.
Bohannon, R. W. (1993). Biomedical applications of hand-held force gauges: A bibliography. *Perceptual and Motor Skills, 77,* 235–242.
Bohannon, R. W. (1995). Internal consistency of dynamometer measurements in healthy subjects and stroke patients. *Perceptual and Motor Skills, 81,* 1113–1114.
Bohannon, R. W. (1997a). Internal consistency of manual muscle testing scores. *Perceptual and Motor Skills, 85,* 736–738.
Bohannon, R. W. (1997b). Reference values for extremity muscle strength obtained by hand-held dynamometry from adults aged 20 to 79 years. *Archives of Physical Medicine and Rehabilitation, 78,* 26–32.
Braud, W. G. (1989). A possible proximity effect on human grip strength: An attempted replication. *Perceptual and Motor Skills, 68,* 157–158.
Brinkmann, J. R. (1994). Comparison of a hand-held and a fixed dynamometer in measuring strength of patients with neuromuscular disease. *Journal of Orthopaedic and Sports Physical Therapy, 19,* 100–104.
Burgess, P. R., & Jones, L. F. (1997). Perceptions of effort and heaviness during fatigue and during the size-weight illusion. *Somatosensory and Motor Research, 14,* 189–202.
Burgess, P. R., & Wei, J. Y. (1984). The effect of ''X-ing'' on competitive motor performance. *Society for Neuroscience Abstracts, 10,* 159.
Byl, N. N., Richards, S., & Asturias, J. (1988). Intrarater and interrater reliability of strength measurements of the biceps and deltoid using a hand held dynamometer. *Journal of Orthopaedic and Sports Physical Therapy, 9,* 399–405.
Caruso, W., & Leisman, G. (2000). A force/displacement analysis of muscle testing. *Perceptual and Motor Skills, 91,* 683–692.
Chorbajian, P. M., Caprarola, M. A., Esterson, P. S., Romero, J. A., Simons, D. A., & Yahara, M. (1988). Muscle strength related to temporomandibular joint function. *Exercise Physiology. Current Selected Research, 3,* 161–169.
Diamond, J. (1979). *Your Body Doesn't Lie. How to Increase Your Life Energy through Behavioral Kinesiology.* New York: Warner Books.
Goodheart, G. J. (1966). Chinese lessons for the modern chiropractic. *Digest of Chiropractic Economics, 8,* 10–15.
Goodheart, G. J. (1976). *Applied Kinesiology* (12th ed.). Detroit, MI: Privately Published.
Hawkins, D. R. (2002). *Power vs. Force. The Hidden Determinants of Human Behavior.* Carlsbad, CA: Hay House.
Hosking, G. P., Bhat, U. S., Dubowitz, V., & Edwards, R. H. T. (1976). Measurements of muscle strength and performance in children with normal and diseased muscles. *Archives of Diseases of Children, 51,* 957–963.
Hsieh, C.-Y., & Phillips, R. B. (1990). Reliability of manual muscle testing with a computerized dynamometer. *Journal of Manipulative and Physiological Therapeutics, 13,* 72–82.
Hyman, R. (1999). Psychology and 'alternative medicine.' The mischief-making of ideomotor action. *The Scientific Review of Alternative Medicine, 3*(2), 34–43.
Jones, L. F., & Burgess, P. R. (1998). Neural gain changes subserving perceptual acuity. *Somatosensory and Motor Research, 15,* 190–199.
Joyce, G. C., Rack, P. M. H., & Westbury, D. R. (1969). The mechanical properties of cat soleus muscle during controlled lengthening and shortening movements. *Journal of Physiology, 204,* 461–474.
Keating, T. M., Kendler, B. S., & Merriman, W. (2004). Evaluation of a possible proximity effect of aspartame and vitamin C on muscular strength. *Perceptual and Motor Skills, 98,* 100–102.
Kendler, B. S., & Keating, T. M. (2003). Re-evaluation of a possible proximity effect on muscular strength. *Perceptual and Motor Skills, 97,* 371–374.
Magnusson, S. P., Gleim, G. W., & Nicholas, J. A. (1990). Subject variability of shoulder abduction strength testing. *American Journal of Sports Medicine, 18,* 349–353.
Monti, D. A., Sinnott, J., Marchese, M., Kunkel, E. J. S., & Greeson, J. M. (1999). Muscle test comparisons of congruent and incongruent self-referential statements. *Perceptual and Motor Skills, 88,* 1019–1028.

Mulroy, S. J., Lassen, K. D., Chambers, S. H., & Perry J. (1997). The ability of male and female clinicians to effectively test knee extension strength using manual muscle testing. *Journal of Orthopaedic and Sports Physical Therapy, 26,* 192–199.

Nicholas, J. A., Sapega, A., Kraus, H., & Webb, J. N. (1978). Factors influencing manual muscle tests in physical therapy. The magnitude and duration of force applied. *Journal of Bone and Joint Surgery, 60-A,* 186–190.

Omura, Y. (1979). Acupuncture (with possible roles of serotonin and melatonin) and related unorthodox methods of diagnosis and treatment: Non-invasive spheno-palatine ganglionic block, abrasion of naso-pharyngeal mucosa, and applied kinesiology. *Acupuncture and Electro-Therapeutics Research, 4,* 69–89.

Perot, C., Meldener, R., & Goubel, F. (1991). Objective measurement of proprioceptive technique consequences on muscular maximal voluntary contraction during manual muscle testing. *Agressologie, 32,* 471–474.

Radin, D. I. (1984). A possible proximity effect on human grip strength. *Perceptual and Motor Skills, 58,* 887–888.

Ryan, A. H., & Agnew, J. H. (1917). Studies in muscular power and fatigue. *American Journal of Physiology, 42,* 599–600.

Sapega, A. A. (1990). Muscle performance evaluation in orthopedic practice. *Journal of Bone and Joint Surgery, 72-A,* 1562–1574.

Scopp, A. L. (1978). An experimental evaluation of kinesiology in allergy and deficiency disease diagnosis. *Journal of Orthomolecular Psychiatry, 7,* 137–138.

Stratford, P. W., & Balsor, B. E. (1994). A comparison of make and break tests using a hand-held dynamometer and the kin-com. *Journal of Orthopaedic and Sports Physical Therapy, 19,* 28–32.

Walther, D. S. (1976). *Applied Kinesiology. The Advanced Approach to Chiropractic.* Pueblo, CO: Systems DC.