ORIGINAL RESEARCH

Emerging from the Mystical: Rethinking Muscle Response Testing as an Ideomotor Effect

Anne M. Jensen, Queensland, Australia

Abstract

Muscle Response Testing (MRT) is an assessment tool estimated to be used by over one million people worldwide, mainly in the field of alternative health care. During a test, a practitioner applies a force on a patient's isometrically contracted muscle for the purpose of gaining information about the patient in order to guide care. The practitioner notes the patient's ability or inability to resist the force and interprets the outcome according to predetermined criteria. Though recent research supports the validity of MRT, little is known about its mechanism of action. Nevertheless, its causation is often attributed to an *ideomotor effect*, which can be defined as muscular activity, potentially nonconscious, and seemingly brought about by a third-party operator. Accordingly, the aim of this study is to investigate whether the ideomotor effect is a plausible explanation of action for MRT.

Methods: This is a retrospective, observational study of data extraction from a previously reported study of the diagnostic accuracy of MRT used to distinguish true from false statements. Additional analysis was carried out on the dataset of assessing for potential sources of bias—both practitioner bias and patient bias.

Results: When practitioners were *blind*, they achieved a mean MRT accuracy of 65.9% (95% CI 62.3–69.5), and when they were *not blind*, 63.2% (95% CI 58.3–68.1). No significant difference was found between these scores (p = 0.37). When practitioners were intermittently *misled,* the mean MRT accuracy decreased to 56.6% (95% CI 49.4–63.8), which proved to be significantly different from when the practitioners were blind (p = 0.02), yet not significantly different from then the practitioners were not blind (p = 0.11). In addition, no evidence of patient bias was uncovered.

Summary: The results of this study demonstrate that when comparing blind and not blind conditions, the practitioner evokes no influence, so it is unlikely that the practitioner is responsible for an ideomotor effect. Likewise, the patient has been shown to produce no significant influence either, so it is also unlikely that the patient is responsible for an ideomotor effect. The limitations of this study are those of any retrospective, observational study in that data were not collected to answer the specific research question of this study. Future research should include a study specifically designed to answer this question, for example, intentionally attempting to induce bias in the practitioner. In summary, the ideomotor explanation of MRT should be regarded as obsolete until such a time as a more plausible explanation of its mechanism of action is established.

Keywords: ideomotor, muscle testing, muscle weakness, lie detection, accuracy, kinesiology

Anne M. Jensen, MSc, DC, DPhil, is a forward-thinking healer who earned her doctorate from Oxford University researching the validity of muscle testing. Through her background in chiropractic and psychology, her empathic ability and sense of curiosity, she developed HeartSpeak (www.heartspeak.me), a unique and empowering stress-reduction tool. **Correspondence:** Anne M. Jensen, 7 Sydney Street, Mackay, Queensland 4740, Australia; email: dranne@drannejensen.com. **Disclosure:** The author receives income from muscle response testing training.

Muscle Response Testing (MRT), commonly referred to simply as "muscle testing," is an assessment tool estimated to be used by over one million people worldwide, mainly in the field of alternative health care, which includes kinesiologists, chiropractors, physiotherapists, osteopaths, and psychologists (Jensen, 2015b). During a test, a practitioner applies a force on a patient's isometrically contracted muscle for the purpose of gaining more information about the patient in order to guide care. The practitioner notes the patient's ability or inability to resist the force and interprets the outcome according to predetermined criteria. Though MRT has been shown to be sufficiently accurate in distinguishing lies from truth (Jensen, Stevens, & Burls, 2016), its mechanism of action is largely unknown, though commonly attributed to an *ideomotor effect*, which can be defined as muscular activity, potentially nonconscious, and seemingly brought about by a third-party operator. The results from a recent study on the accuracy of MRT suggest, however, that the ideomotor effect may not be the mechanism of action. Therefore, the purpose of this paper is to define what MRT is, then to review the evolution and key features of the ideomotor effect, and finally to outline the specific results of the recent study that preclude an ideomotor effect as a plausible explanation of MRT's mechanism of action.

Muscle Response Testing

MRT is one type of manual muscle testing (MMT) in which a patient's muscle—often the deltoid muscle—is tested repeatedly as the target condition changes. Conventionally, the practitioner detects the outcome of the muscle test to be either "strong" or "weak." The interpretation of the outcome (strong or weak) is dependent upon the practitioner's choice of paradigm. A common usage of MRT is to distinguish true from false spoken statements, and a common paradigm employed is: A true statement results in a strong muscle response, and a false statement results in a weak muscle response (Jensen et al., 2016; Monti, Sinnott, Marchese, Kunkel, & Greeson, 1999). Another common usage of MRT is to detect stress in a patient (Gallo, 2000; Frost & Goodheart, 2013; Thie & Thie, 2005; Krebs & McGowan, 2013), and in this paradigm, a weak muscle response is usually indicative of stress, and a strong response usually suggests the absence of stress.

MRT differs distinctly from other forms of MMT, such as orthopedic/neurological MMT (O/N-MMT) and Applied Kinesiology–style MMT (AK-MMT). In both of these other forms, any muscle of the body may be tested, whereas in MRT, only one muscle is usually tested and it is commonly called the *indicator muscle*. In O/N-MMT, muscles are tested to assess muscle strength and graded on a 0–5 scale, with 5 being normal. In MRT (and in AK-MMT), muscles are tested to assess conditions other than strength and are graded on a binary scale: either strong or weak. Another distinction between AK-MMT and MRT is that in AK-MMT, the test outcome is dependent upon the muscle being tested. For instance, if a psoas muscle is found to be weak, this may indicate a kidney concern (Walther, 1981, 2000). Whereas in MRT, the practitioner decides on the parameters of the test prior to its execution. As an example, a patient may report having neck pain, the practitioner may perform MRT using the deltoid muscle to detect the presence (or absence) of stress in the neck, and a weak response will indicate that stress is present. The practitioner may then continue to use MRT to zero in on the source of the stress. For a summary of the similarities and differences of the three types of MMT, see Table 1.

The Current Status of the Evidence of MRT

Until the development of the Standards for the Reporting of Diagnostic Accuracy (STARD) guidelines in 2003, the evaluation of diagnostic techniques lagged behind that of interventions and had

Table 1. *The Three Types of Manual Muscle Testing (MMT): A Summary of Their Similarities and Differences*

	O/N-MMT	AK-MMT	MRT
Muscles used	Any muscle	Any muscle	One muscle*
Detects	NMS conditions	Many conditions	Many conditions
Outcomes	Graded 0 to 5	Binary	Binary
Results	Assesses strength	Depends on muscle tested	Depends on condition

Note: *An Indicator Muscle.
Abbreviations: O/N-MMT = Orthopedic/Neurological Manual Muscle Testing; AK-MMT = Applied Kinesiology–style of Manual Muscle Testing; MRT = Muscle Response Testing; NMS = Neuromusculoskeletal.

been notoriously fraught with inconsistencies and bias (Knottnerus & Buntinx, 2009; Bossuyt et al., 2003b; Hall, Lewith, Brien, & Little, 2008). This is especially true of the inconsistent use of terminology to describe the validity of a diagnostic test: Various terms (e.g., accuracy and precision) are confused in colloquial English and at times in the scientific literature as well (Slezák & Waczulíková, 2011). In assessing the current status of the MRT literature, this difficulty is further amplified by the confusion regarding the term "muscle testing," which, as previously described, can have different meanings in different contexts.

With this in mind, using the electronic databases MEDLINE, MANTIS, PsycINFO, and CINAHL, a literature search was conducted, and only papers published in peer-reviewed journals were considered. The outcome of this search was 26 papers that used either MRT or AK-MMT to detect a specified target condition. The reference lists of the included papers were also checked for relevant research, which resulted in no additions.

Few rigorous studies have attempted to estimate the diagnostic accuracy of MRT or AK-MMT; however, one must take into account that only one of these, Schwartz et al. (2014), was published after the publication of the STARD guidelines. This study attempted to measure the diagnostic accuracy of MRT to distinguish between substances that were toxic and nontoxic to the body, and reported that MRT accuracy was indistinguishable from chance. This means, in this case, that MRT was unsuccessful. Likewise, two systematic reviews of the AK literature (Hall et al., 2008; Klinkoski & Leboeuf, 1990) found no evidence of diagnostic accuracy, but the standards for reporting were pre-STARD, and therefore lacking.

There are, however, numerous studies that have looked at other characteristics of MMT, such as reliability (Drouin, Valovich-McLeod, Shultz, Gansneder, & Perrin, 2004; Florence et al., 1992; Haas, Peterson, Hoyer, & Ross, 1994; Jepsen, Laursen, Larsen, & Hagert, 2004; Perry, Weiss, Burnfield, Gronley, 2004; Pollard, Lakay, Tucker, Watson, & Bablis, 2005; Wadsworth, Krishnan, & Sear, 1987), validity (Drouin et al., 2004; Perry et al., 2004; Ladeira et al., 2005), inter-examiner agreement (Pollard et al., 2005; Lawson & Calderon, 1997), intra-examiner agreement (Wadsworth et al., 1987; Leboeuf, Jenkins, & Smyth, 1988), predictability (Pollard, Bablis, & Bonello, 2006; Perry, Ireland, Gronley, & Hoffer, 1986), internal consistency (Bohannon, 1997), and diagnosis in general (Jacobs, Franks, & Gilman, 1984; Nahmani, Serviere, & Dubois, 1984; Omura, 1981; Pothmann, Hoicke, Weingarten, & Lüdtke, 2001; Schmitt & Leisman, 1998; Tiekert, 1981; Triano, 1982). The sheer number of terms used to describe the validity of MRT is frankly confusing. Moreover, the appropriateness of the application of some of these analyses to MRT or AK-MMT is questionable. However, some published studies do report accuracy estimations, most of them published prior to the publication of the STARD guidelines.

Using the AK-style of MMT, Caruso and Leisman (2000) reported that experienced practitioners (≥5 years' experience) predicted muscle strength more accurately compared to inexperienced practitioners (<5 years' experience), with accuracies of 98% and 64%, respectively. In other studies, it was found that MRT was used to accurately predict low back pain (Pollard et al., 2006) and simple phobia (Peterson, 1996), and AK-MMT accurately predicted food allergies (Garrow, 1998). Further studies found that AK-MMT was unable to accurately predict nutritional needs (Triano, 1982; Kenney, Clemens, & Forsythe, 1988), nutritional intolerance (Pothmann et al., 2001; Jacobs, 1981), and thyroid dysfunction (Jacobs et al., 1984).

Nevertheless, one study, Monti et al. (1999), successfully used MRT to differentiate between true and false statements, similar to the current study (Jensen et al., 2016). Monti et al. found that when a muscle is tested following a true spoken statement, it yields significantly different results compared to MRT following false spoken statements. Their study found that the indicator muscle stays "strong" after a patient speaks true statements and goes "weak" after a patient speaks false statements. Their statements were self-referential statements, which used the speaker's name, as in "My name is (insert one's name or another name)." One problem with using self-referential statements is that, in all likelihood, both the muscle tester and the test patient were aware of the verity of the statement, and therefore neither was blind. Blinding of the muscle testers was not specifically reported in Monti's paper. Even though tester bias was examined in Monti's study, there is a chance that unblinded testing may have introduced other biases and thus influenced the test's outcome. While it is generally accepted among those who use various types of MMT that some bias can exist, little is currently known about the degree of this bias.

Development of Ideomotor Principles

The term *ideomotor* is derived from *ideo*, meaning "an idea or mental construct," and *motor*, meaning "muscular activity," suggesting that muscle movement can be driven by thoughts (Stock & Stock, 2004). While the phrase "ideomotor effect" was coined by William Carpenter in 1852, its principles stem from earlier that century and have two clear roots: one British and the other German. The main difference between the two theories is that in the German model (the older model), the ideomotor effect can be brought about by either conscious (voluntary) or nonconscious (involuntary) effort, while the British primarily focused on nonconscious muscular activity (Stock & Stock, 2004; Braid, 1855). An evolutionary tree outlining the primary influencers of ideomotor theory can be found in Figure 1.

Carpenter's predecessor, Thomas Laylock, observed in 1845 that people appeared at times to have "no deliberate control over their own behavior," and also acknowledged this was not a new concept but dated back to the 17th and 18th centuries (Stock & Stock, 2004). Another influence on ideomotor principles was James Braid, the founder of modern hypnotism and Carpenter's good friend and sounding board. Braid used Carpenter's new ideomotor effect to explain some of the phenomenon he observed in hypnotized subjects. For example, a hypnotherapist (also called the "operator") may suggest to the hypnotized subject that he cannot rise from the chair or open his eyes, so the subject takes on this belief and, as a result, finds that it is actually impossible for him to rise or open his eyes—his muscles will not allow it. Braid and Carpenter suggested that this is not because the will of the subject is controlled by the operator but rather that his will is controlled by his own belief in the operator's suggestion (Carpenter, 1852)—an important but subtle distinction.

Braid and Carpenter extended this principle beyond motor/muscular control and suggested that any reflex action can be elicited through thought. For instance, at the suggestion that the room is exceptionally hot, a subject may actually perspire (Carpenter, 1852). As a result of its connection with hypnosis, the ideomotor effect then came to be singularly associated with a seeming loss of voluntary control, whereas the original theory was concerned with ideas exerting both voluntary and involuntary control over muscular actions (Braid, 1855).

At the same time that Braid formalized hypnotherapy and Carpenter proposed his ideomotor theory, occultism was on the rise. There was widespread use of the pendulum, divining rod, and Ouija board, and seances and mediumship were popular. Largely due to proximity, and also the lack of other plausible explanations, many of these paranormal phenomena were explained by

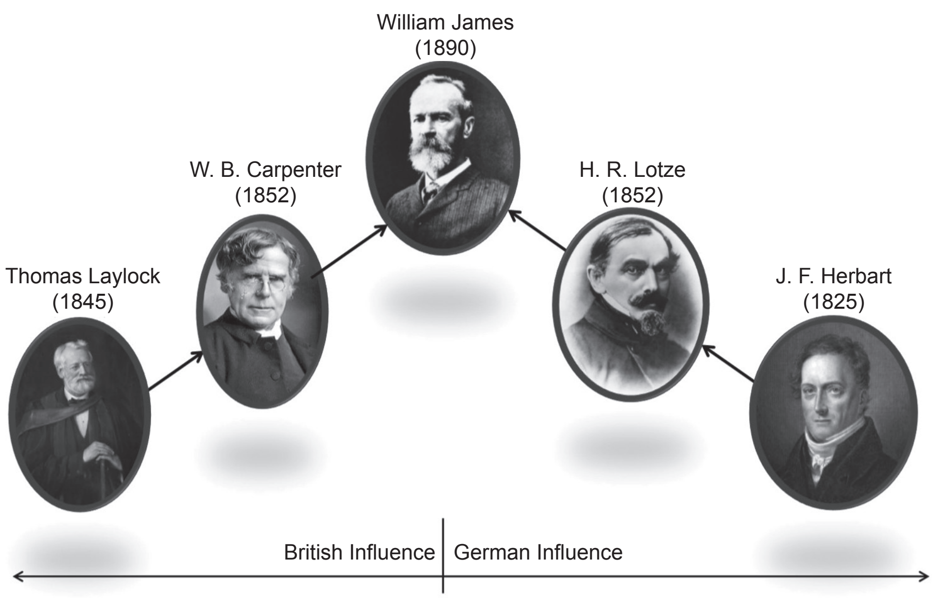

Figure 1. Influencers of the ideomotor theory.

ideomotor action. Coincidentally, because of the popularity of and the fascination with the occult that enchanted Europe in the mid-19th century, Carpenter's ideomotor effect also gained widespread support (Stock & Stock, 2004). Consequently, by the end of the 19th century, the two—the ideomotor effect and the occult—became intertwined in the public mind and that connection persists to this day.

William James (1890s), often regarded as the Father of American Psychology, was committed to demystifying the mechanisms of movement and discovering their true neurocognitive pathways (Ondobaka & Bekkering, 2012). In attempt to disentangle the ideomotor effect and the occult and add credibility to the emerging discipline of modern psychology, James adopted Carpenter's term *ideomotor* but opposed its application to occult or paranormal phenomena (Stock & Stock, 2004). Rather, James suggested that any action may be the result of an ideomotor effect; thus, essentially, he adopted the British term but the German principles (Stock & Stock, 2004).

In an attempt to put more distance between the modern psychological principles and the occult, Edward Lee Thorndike, then president of the American Psychological Association, vehemently and publicly attacked ideomotor theory in his speech at their 1912 conference (Stock & Stock, 2004). As a result of Thorndike's scathing rebuke, the original ideomotor principles fell into oblivion, and to this day the term "ideomotor effect" is attached to paranormal phenomena and is looked upon with repugnance.

It is in the category of ideomotor effect that muscle testing has been incorrectly placed today, alongside the pendulum, dowsing, Facilitated Communication, automatic writing, the Ouija board, and other unexplained phenomena. Muscle testing's poor face validity and the lack of rigorous evidence of its mechanism of effect may have contributed to this continued categorization. Despite its lack of evidence of effect, however, the mechanism cannot be justified by the ideomotor effect.

Defining the Ideomotor Effect

As the meaning of *ideomotor* has changed markedly since its inception, for the purpose of clarity, this paper will use the following contemporary definition:

> *Ideomotor effect*: the process whereby a thought brings about a seemingly reflexive or automatic muscular action, potentially slight, and conceivably outside the subject's awareness (Shin, Proctor, & Capaldi, 2010).

In light of this definition and upon review of the literature, there seem to be two consistent features of the ideomotor effect (Stock & Stock, 2004; Ondobaka & Bekkering, 2012; Carpenter, 1852, pp. 147–153; Carroll, 2012; Hyman, 1999; Jackson, 2005):

1. The subject seems to have no control over his/her actions; and
2. Ideas seem to be at the suggestion of the operator.

Putting these features in terms of MRT, the translation might be as follows:

1. The patient seems to have no control over the strength of his/her muscle; and
2. The practitioner may be suggesting to the patient the outcome desired, or in other words, the practitioner seems to exert an influence during the test, or in still other words, seems to bias the muscle test.

With these features in mind, it is easy to understand how the ideomotor effect was used as a plausible explanation of MRT: A patient's muscle seems to stay strong during one test and then goes weak during the very next test, with the patient seeming to have no control.

If data are uncovered to disprove either of these points, however, then MRT cannot be credited to an ideomotor effect. This study will assess the second feature and investigate if blinding the participants influences MRT accuracy. It is reasonable to theorize if the practitioner is not blind to the expected outcome of the MRT (i.e., s/he knows what the outcome should be), then s/he will or can introduce bias and try to sway the test toward the expected outcome, thereby artificially inflating the accuracy. Consequently, the null hypothesis (H_0) of this study is that MRT accuracy when the practitioner is not blind is greater than MRT accuracy when the practitioner is blind:

H_0: MRT accuracy (not blind) > MRT accuracy (blind)

If blinding has no effect on MRT accuracy, then the null hypothesis will be rejected, demonstrating that MRT cannot be explained by an ideomotor effect.

Methods

Using the results from a previously reported study on the accuracy of MRT (Jensen et al., 2016), additional analyses were made on the dataset of Study 1, focusing on the second of the two ideomotor features described above: the influence of bias during a muscle test.

The study previously reported and used for additional analysis was a prospective study of diagnostic test accuracy, registered with two clinical trials registries (the Australian New Zealand Clinical Trials Registry [ANZCTR; www.anzctr.org.au; ID#ACTRN12609000455268] and the US-based ClinicalTrials.gov [ID#NCT01066312]), and received ethics committee approval to collect data in the United Kingdom and the United States. Data collection in the United Kingdom was approved by the Oxford Tropical Research Ethics Committee (OxTREC Reference Number 34-09), and data collection in the United States, by the Parker University Institutional Review Board (Approval Number R09-09). The study under further analysis was the first in a series of studies undertaken for the degree of DPhil (PhD), which was granted by Oxford University in 2015 (Jensen, 2015a). Written informed consent was obtained from all participants, and all other tenets of the Declaration of Helsinki were upheld. In addition, this study was reported in accordance with the Standards for the Reporting of Diagnostic Test Accuracy Studies (STARD) guidelines (Bossuyt et al., 2003a; Bossuyt et al., 2003b; Bossuyt & Leeflang, 2008).

General Summary of Study Methods

In this study, 48 practitioners experienced in MRT (Practitioners) were paired with 48 naïve and unique test patients (TPs), who had no prior experience with MRT. Both types of participants, Practitioner and TP, each viewed their own computer screen, on which pictures of common objects were displayed. TPs were instructed via an earpiece to speak a statement in reference to the displayed picture; sometimes the statement was true, other times it was false. (For a pictorial description of the testing scenario layout, see Figure 2.) The sequence of true and false statements was randomly generated by the computer. In Part 1 of this study, Practitioners were shown either the same picture as the TP or a blank, black screen. In the latter case, the Practitioner was considered *blind*, and in the former, *not blind*. The Practitioner-blind condition was used to calculate mean MRT accuracy (as percent correct). The Practitioner-not blind condition was used to assess if practitioners could be persuaded (or biased) toward choosing the correct response.

Once the TP viewed the picture and spoke the given statement, the Practitioner performed a deltoid muscle test in order to determine if the statement was true or false. At this point, the TP entered the MRT outcome on his/her computer keyboard, which advanced the screens of TP and Practitioner to the next picture/statement. Part 1 consisted of 40 MRTs, and Part 2, of 20 MRTs. In Part 2, the Practitioner's computer also randomly displayed pictures that were different from the TP's, creating the "Practitioner-misled" condition, and potentially a bias away from choosing the correct response. See Figure 2 for a diagram of the testing scenario, and Figure 3 for the Participant Flow Diagram. (For a more detailed description of these study methods, refer to the original paper [Jensen et al., 2016].)

Blinding of Participants

Since it was an aim of the original study to investigate the impact that blinding had on MRT accuracy, much thought was given about how to blind participants, both TPs and Practitioners. Blinding of participants is the key feature of this report.

First, since TPs were MRT-naïve, they were unaware of what MRT involved and were unfamiliar with any MRT paradigms. Also, while it was impossible to blind TPs to the reaction of their arms, they were not explicitly told when their arm stayed strong or went weak, and in many instances, the Practitioner's determination of test outcome ("strong" or "weak") was not obvious to myself, an observer, during testing. In addition, because Practitioners were randomly blinded, TPs were effectively blind to the Practitioner's blindness. Furthermore, TPs were blind to the interpretation of the MRT outcome ("Weak" was interpreted as "lying" and "Strong" was interpreted as "truth") and blind to what the Practitioner entered into the computer (e.g., "W" or "S"). Furthermore,

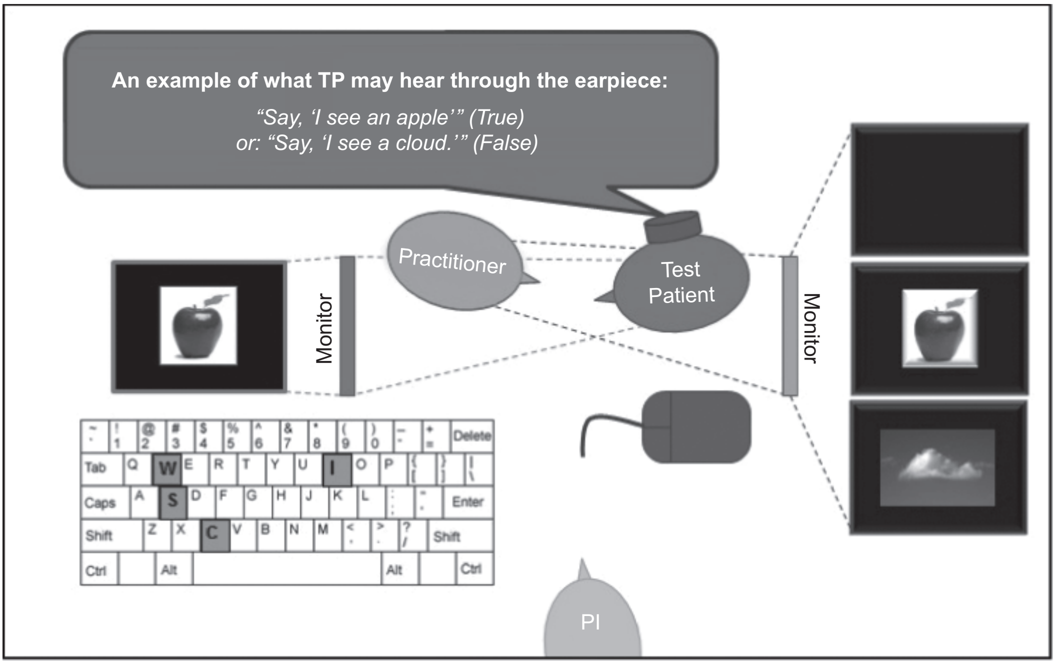

Figure 2. Testing scenario layout: The Practitioner (light gray) viewed a monitor (also light gray) that the Test Patient (TP) could not see and entered his results on a keyboard. The Test Patient (dark gray) viewed a monitor (also dark gray) that the Practitioner could see, had an ear piece in his ear through which he received instructions, and used a mouse to advance his computer to the next picture/statement. Note that in the part of this study currently under analysis, Practitioners were presented with (1) the same picture as the Test Patient, (2) a blank, black screen, or (3) a picture that was different from the Test Patient's picture. Also note that the Principal Investigator (PI) was present in the room and observing during all assessments.
Abbreviations: TP = Test Patient; PI = Principal Investigator.

no findings or results were discussed with the TP during the testing, nor was the TP's opinion sought. Finally, TPs were theoretically not blind to the verity of the statements they spoke. That is, it was presumed that TPs were aware of when they spoke true statements and when they spoke false statements.

On the other hand, blinding the Practitioner was, in many respects, more straightforward. To begin with, clearly Practitioners were *not* blind to the paradigm being used. They were also aware of: (1) the primary aim of the study (i.e., to estimate MRT accuracy in detecting deceit), (2) that TPs were naïve to MRT, (3) that TPs were unaware of the paradigm being used, and (4) TPs were going to be instructed either to lie or to tell the truth. Practitioners were, however, randomly blind to the verity of the spoken statement. In addition, in a second part of the study (Part B), when they were misled about the sameness of the picture they were shown, they were also blind to being blind. Taking all these factors into consideration, I believe that overall a high level of participant blinding was achieved.

Data Extraction Methods

Specific data were extracted from the study's dataset to assess for participant bias: Practitioner bias and TP bias. Practitioner bias was evaluated by comparing the mean MRT accuracies under three different conditions: (1) when the Practitioner was *blind*, (2) when the Practitioner was *not blind*, and (3) when the Practitioner was *misled*. While the primary outcome of this study was the mean MRT accuracy when the Practitioners were blind, if this measure was significantly different from either of the other two conditions, then it might suggest bias was present.

TP bias was assessed by subgroup analysis by comparing the mean MRT accuracy of the group where TPs reported guessing the paradigm being investigated to the mean MRT accuracy of the group who did not report guessing the paradigm

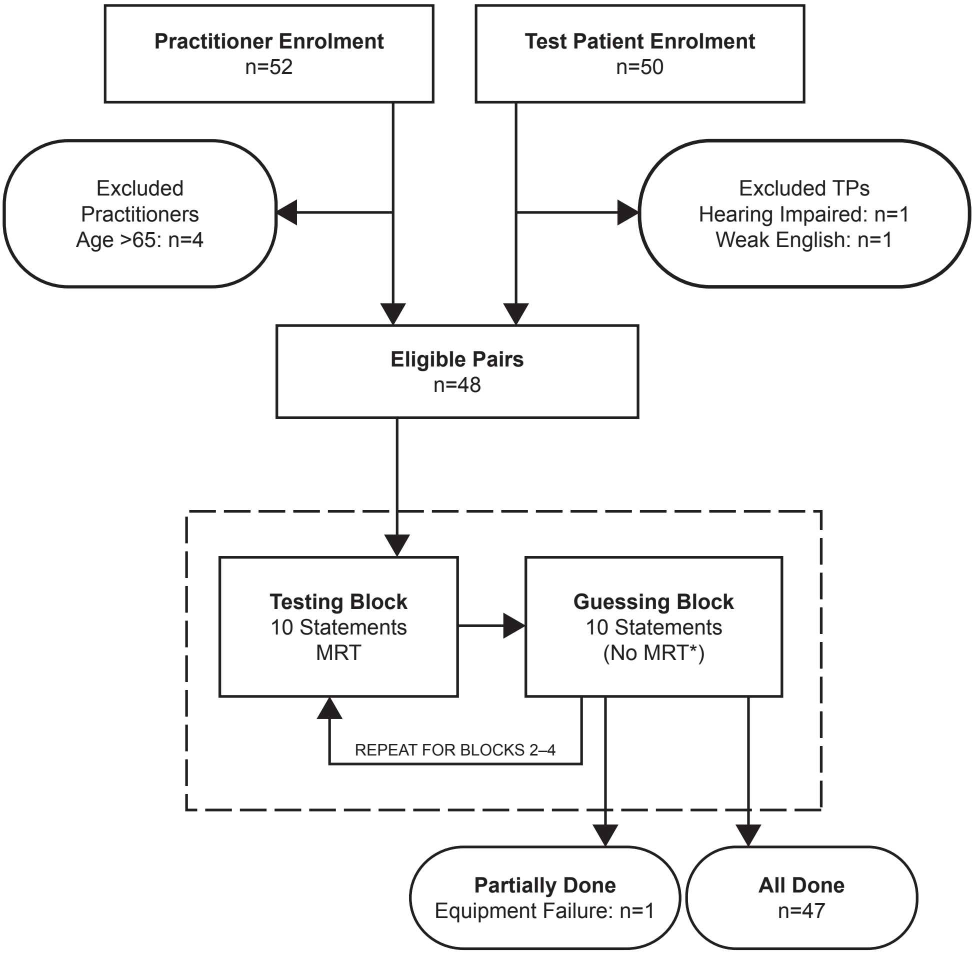

Figure 3. Participant flow diagram.
Note: *Touching wrist & observing.
Abbreviations: MRT = Muscle Response Testing; TP = Test Patient.

under investigation. It was hypothesized that if a TP was *not blind* to the paradigm under investigation, then s/he could introduce a degree of response bias.

Finally, some researchers have noted a heightened belief in a process following a positive personal experience—a phenomenon they (possibly inaccurately) labeled an "ideomotor action" (Hyman, 1999; Dillon, Fenlason, & Vogel, 1994). In any event, in the original study, participants were asked to rate their levels of confidence on a 10 cm Visual Analog Scale (VAS) before and after testing. TPs ranked the level of confidence they had in MRT in general, in their Practitioner, and in their Practitioner's MRT ability. Similarly, Practitioners rated their level of confidence in MRT in general and in their own MRT ability. While not proven to be correlated to positive personal experience, confidence rating data were reanalyzed and re-reported.

Due to word limit restraints imposed by scientific journals, some of the results of the original study presented here have not been previously reported yet were peer-reviewed during the degree conferring process (Jensen, 2015a).

Statistical Methods

An error-based measure of MRT accuracy is reported as overall percent correct, with a 95% confidence interval (95% CI). Using pilot data, a sample size for this full-scale study was calculated, powered to 80%. Based on these assumptions and using a 95% confidence interval, it was determined that a study of 48 practitioner-patient pairs would have good statistical power to demonstrate whether MRT can be used to distinguish a lie from a truth. Statistical advice was sought during the design phase, after piloting, and before data analysis. All data were analyzed using Stata/IC 12.1 (StataCorp LP, College Station, Texas), specifically the commands *ttest* and *pwcorr*.

Results

Specific data were extracted from the dataset of Study 1 in a previously reported series of diagnostic test accuracy studies on MRT. (For the results of this series, see Jensen et al., 2016).

Bias Toward a Desired Outcome: Blind vs. Not-Blind Practitioners

In Part 1 of this study, MRT accuracies were compared when the Practitioners were blind to when they were not blind. When the Practitioners were blind, they achieved a mean MRT accuracy of 65.9% (95% CI 62.3–69.5), and when they were not blind, 63.2% (95% CI 58.3–68.1). See Table 2. No significant difference was found between the scores ($p = 0.37$), and they were significantly correlated ($r = 0.383$, $p = 0.01$).

Bias Away from a Desired Outcome: An Attempt to Mislead Practitioners

In this previously unreported part of this study, where Practitioners were intermittently *misled*, mean MRT accuracy dropped to 56.6% (95% CI 49.4–63.8), which proved to be significantly different from when the Practitioners were *blind* ($p = 0.02$), yet not significantly different from then the Practitioners were *not blind* ($p = 0.11$). See Table 2.

Table 2. *The Impact of Blinding and Misleading Practitioners on MRT Accuracy*

Condition		MRT accuracy†		
		Mean	**95% CI**	***p*-value**
Part 1	Blind	0.659	0.623–0.695	0.37
	Not blind	0.632	0.583–0.681	
Part 2	Not blind	0.639	0.595–0.693	} 0.53
	Blind	0.659	0.623–0.695	
	Misled	0.566	0.494–0.638	} 0.02*

Notes: †Accuracy = % correct.*Significance reached.
Abbreviations: MRT = Muscle Response Testing; CI = Confidence Interval.

Test Patient Bias

Also compared were the mean accuracies of those pairs whose TP reported guessing the paradigm ($n = 21$) to those pairs whose TPs did not report guessing the paradigm ($n = 27$). For those pairs whose TP reported guessing the paradigm, the mean MRT accuracy was 66.1% (95% CI 59.1–73.0), and for those pairs whose TP did not report guessing the paradigm, the mean MRT accuracy was 64.9% (95% CI 61.0–68.8). No significant difference was found between these two groups ($p = 0.38$).

Pre- and Posttesting Confidence Ratings

Overall, while TP confidence ratings increased significantly (see Table 3A), this did not seem to influence MRT accuracy (see Table 3B). Furthermore, while Practitioner confidence ratings dropped slightly, their differences did not reach significance (see Table 3A). Moreover, these ratings were not correlated with MRT accuracy scores, regardless whether practitioners were *blind* or *misled* (see Table 3C). However, MRT accuracies under blind conditions did correlate with MRT accuracies in the misled condition.

Discussion

In this further analysis of the dataset from a previously reported study, sources of potential bias were assessed. In Part 1 of this study, Practitioners were given an opportunity to introduce bias into their testing by being given suggestions about the verity of the TP statements (i.e., they were intermittently made aware when a statement was true or false). For this part, the mean MRT accuracy when the Practitioners were *blind* was compared to when they were *not blind*, and the difference was found to be insignificant ($p = 0.37$). This suggests that even when the Practitioners knew what the outcome of the muscle test should be, they performed similarly to when they did not know, suggesting that they did not exert an influence on (i.e., bias) the outcome of the test.

In Part 2 of this study, Practitioners were given further opportunities to induce bias into their muscle testing when it was attempted to sway them away from—and also again toward—choosing the correct response. When the MRT accuracy when the Practitioners were *blind* (65.9%; 95% CI 62.3–69.5) was compared to the MRT accuracy

Table 3A. *Comparing Pre- and Posttesting Confidence Ratings*

		Pretesting		Posttesting		*p*-value
		Mean rating	95% CI	Mean rating	95% CI	
Test patients	Confidence in MRT in general	6.76	6.16–7.36	7.22	6.63–7.81	0.03*
	Confidence in Practitioner	6.95	6.30–7.61	7.63	7.01–8.25	0.01*
	Confidence in Practitioner's MRT ability	7.00	6.35–7.65	7.76	7.10–8.41	0.01*
Practitioners	Confidence in their own MRT ability	8.43	8.02–8.85	8.15	7.67–8.63	0.37
	Confidence in MRT in general	8.67	8.22–9.12	8.43	7.94–8.92	0.47

Note: *Significance reached.
Abbreviations: MRT = Muscle Response Testing; TP = Test Patient; CI = Confidence Interval.

when the Practitioners were misled (56.6; 95% CI 49.4–63.8), a significant difference was found ($p = 0.02$), suggesting that, in this case, the Practitioners exerted an influence. On the other hand, when comparing *blind* vs. *not blind* Practitioners' MRT accuracies in Part 2 (as was done in Part 1), the difference in the means was likewise found to be insignificant ($p = 0.53$). This once again suggests that Practitioners did not exert an influence.

Potential sources of TP bias were also evaluated. Since in this testing scenario, the TPs were not themselves blind to the verity of their spoken statements, it was possible for them to exert an influence on the MRT outcome. However, only naïve volunteers were recruited into this study as TPs. They had no prior experience with MRT, and each TP also did not know the Practitioner who was performing the testing. In addition, TPs were kept blind to the paradigm being used; that is, they were not explicitly told that a true statement would result in a strong MRT and a false statement would result in a weak MRT. As anticipated, some TPs reported guessing this paradigm; however, no significant difference between these groups was found ($p = 0.38$).

Previous research reported that a positive personal experience could heighten belief in a process, which they (perhaps inappropriately) called an "ideomotor action" (Hyman, 1999; Dillon et al., 1994). While this current study did not track on "positive-ness" of experience, participants were asked to provide confidence ratings. However, no correlation was found between MRT accuracy and any confidence rating.

Possible Explanations and Implications in Regard to the Ideomotor Effect

The fact that practitioners performed similarly whether they were *blind* or *not blind* suggests that they were not exerting an influence on MRT outcomes. Correspondingly, this makes sense: If an operator is blind, then s/he cannot suggest a correct outcome. This, then, contradicts the

Table 3B. *Correlation Table (including p-values): MRT Accuracy and the Difference in Test Patient Confidence Ratings (r)*

		1.	2.	3.	4.
1. The change in TP confidence in MRT in general: Post–Pre		1.00			
2. The change in TP confidence in Practitioner: Post–Pre		0.33	1.00		
	p-value:	0.02*			
3. The change in TP confidence in Practitioner's MRT ability: Post–Pre		0.30	0.87	1.00	
	p-values:	0.04*	<0.01*		
4. MRT accuracy		0.12	0.06	0.09	1.00
	p-values:	0.41	0.71	0.56	

Note: *Significance reached.
Abbreviation: MRT = Muscle Response Testing.

Table 3C. *Correlation Table (including p-values): MRT Accuracy and the Difference in Practitioner Confidence Ratings (r)*

		1.	2.	3.	4.
1. Change in Practitioner's confidence in MRT in general: Post–Pre		1.00			
2. Change in Practitioner's confidence in own MRT ability: Post–Pre		0.50	1.00		
	p-value:	<0.01*			
3. MRT accuracy (blind condition)		0.14	0.06	1.00	
	p-values:	0.34	0.67		
4. MRT accuracy (misled condition)		0.04	0.06	0.35	1.00
	p-values:	0.79	0.68	0.02*	

Note: *Significance reached.
Abbreviation: MRT = Muscle Response Testing;

second feature of an ideomotor effect, in that the subject is influenced by an idea suggested by the operator. Consequently, in this regard, MRT cannot be explained by an ideomotor effect.

On the other hand, during the condition when the Practitioners were *misled*, MRT accuracy diminished (compared to the *blind* condition), which seems to imply that Practitioners did exert a bias in Part 2. This may be one explanation of these results; there may be other explanations. For example, since Practitioners were required to perform 60 MRTs, fatigue may have contributed to their underperformance in Part 2. However, since there was no significant difference between the MRT accuracy in the *blind* condition of Part 2 and the MRT accuracy in the *blind* condition of Part 1, fatigue is an unlikely explanation. Alternatively, the drop in MRT accuracy in the *misled* condition of Part 2 could mean that Practitioners started to doubt themselves while being misled. This parallels their slight drop in confidence ratings (from pre- to posttesting), a change that may have reached significance had this study been powered for subgroup analysis.

Similarly, it is also unlikely that TPs exert an ideomotor effect on MRT outcomes. This is supported by the result that whether or not TPs reported guessing the paradigm, MRT accuracy scores were not impacted. Likewise, because no correlation was found between MRT accuracy and any confidence rating, this further supports the argument that MRT cannot be explained by an ideomotor effect.

Another explanation of these results may explain a source of confusion. It may be that the terminology in use today is either inaccurate or inconsistent. Perhaps the way the term *ideomotor effect* is currently used does not match the way it was and is defined. For example, in neuro-linguistic programming (NLP), an ideomotor response is defined as "unconscious physical manifestations of mental events" (Wingett, 2013). Examining MRT in regard to this definition and the first feature described previously (i.e., the subject seems to have no control over his/her actions), then the ideomotor effect may be a plausible explanation of MRT. However, this NLP definition does not incorporate the second feature of an ideomotor response, that of the operator's (practitioner's) influence (or bias). Since the study described here suggests that practitioners do not seem to bias the MRT, there may be two effects at play. If this is the case, then perhaps another term is needed for the NLP phenomenon, because the use of the term *ideomotor effect* is not accurate in this case.

Strengths and Limitations

Clear strengths of the original study include a high degree of blinding and the choices of reference standard and target condition. It is commonly thought that practitioners can introduce a great deal of bias during MRT. One way to limit practitioner bias was accomplished through random blinding of Practitioners. Similarly, it is thought that patients can introduce bias in MRT by letting their arms go weak at will, an example of response bias (McGrath, Mitchell, Kim, & Hough, 2010; Haas et al., 1994) or social desirability bias (King & Bruner, 2000); however, much effort was made to keep TPs blind as well. Furthermore, the choices of reference standard and target condition were clear and well defined. In studies of diagnostic test accuracy, it is presumed that the target condition is either present or absent (Martin & Lovett, 1915) and, ideally, the best available method for detecting the presence or absence of the target condition is used as the reference standard (Bossuyt et al., 2003b). Ideally, studies of diagnostic test accuracy employ a true "gold" standard, which demonstrates perfect accuracy in distinguishing the presence or absence of the target condition. However, perfect gold standards are rare in medical testing, so an imperfect reference standard is normally employed (Glasziou, Irwig, & Deeks, 2008). In the original study, the reference standard was the actual verity of the spoken statements, each of which was definitively known to be either true or false—a perfect reference standard; it may therefore be considered a true "gold" standard.

A limitation of the original study is its lack of generalizability to other applications of MRT. While MRT may be useful in distinguishing false statements from true statements, this study does not determine whether MRT is useful for other applications, such as detecting a food allergy (Garrow, 1998; Teuber & Porch-Curren, 2003) or the need for homeopathy (Moncayo, Moncayo, Ulmer, & Kainz, 2004) or a nutritional supplement (Triano, 1982). This point is important to emphasize due to the widespread and varied use of MRT.

Limitations of this present study include those regarding its retrospective, observational design, drawing on data previously presented. For instance, data were not collected to answer the specific research question of this study, and the use of secondary data of this nature may be criticized (Tripathy, 2013; Emma, 2008). For example, other biases may have been introduced, such as misclassification bias, missing variables (unmeasured confounding), and missing data (Benchimol et al., 2015).

In this study, no mechanical testing devices were utilized, such as force plates or a dynamometer. Some may consider this a limitation of this study; however, this point was considered carefully and rejected. This study attempted to reproduce a real clinical setting, and force plates are not routinely used in clinical practice. Supporting this decision, previous studies using force plates showed a distinct difference between muscles labeled "strong" and "weak" (Monti et al., 1999; Caruso & Leisman, 2001; Conable, Corneal, Hambrick, Marquina, & Zhang, 2006), making their use in this study redundant.

Only two other published papers were found in the scientific literature that mention the ideomotor effect in regard to MRT. One study was a review of the literature and a commentary (Schmitt & Cuthbert, 2008), and because it is not experimental in design, it will not be included in this discussion. In a second study, Pollard et al. (2011) report assessing the reliability of MRT to detect low back pain, while also attempting to minimize any potential impact of an ideomotor effect, which they defined as "the unconscious and inadvertent cueing of desired responses." They reported accomplishing this by minimizing all nonverbal and visual cues, especially in regard to pain status, and by eliminating any other communication between practitioner and patient. While great efforts were clearly made to minimize inadvertent cueing, it is unclear if this was successfully accomplished since no assessments of these efforts were reported. Furthermore, their definition of ideomotor effect may not have been in alignment with current consensus (Stock & Stock, 2004; Ondobaka & Bekkering, 2012; Shin et al., 2010).

Directions for Future Research

While the results of this study cast doubt on whether MRT can be explained by the ideomotor effect, more research is certainly required. In regard to MRT, the two features identified as characterizing an ideomotor effect are as follows:

1. The patient seems to have no control over the strength of his/her muscle; and
2. The practitioner may be suggesting to the patient the outcome desired, or in other words, the practitioner seems to exert an influence during the test, or in still other words, seems to bias the muscle test.

The results of this study challenge the second feature. While on the one hand, clinically it makes sense that the practitioner may indeed have an influence on MRT outcomes, in this study when *blind* and *not blind* conditions were compared, no practitioner influence was observed. Further prospective research is needed to assess and, if possible, quantify and qualify the type of influence a practitioner may evoke.

Likewise, future researchers may also wish to investigate the first feature of an ideomotor effect, by answering the question: *In MRT, does the subject have any control over his/her actions?* To assess this, studies must be designed to try to influence (or bias) the patient. In addition, this may also involve finding ways to successfully blind the patient during MRT.

Should future research support the findings of this study, and it become established that MRT cannot be explained by the ideomotor effect, then future researchers may wish to investigate other mechanisms of actions. There are many theories used to explain the weakening of a muscle during a muscle test, such as: (1) a stress (Jensen, 2012), (2) a disruption in homeostasis (Thie & Thie, 2005; Walther, 1988), and (3) other similar disturbances (Rolfes, 1997). Since none of these hypotheses have been substantiated by rigorous science, future research may wish to investigate the plausibility of these and other theories that may arise.

On the other hand, I believe that there are much more important research questions to answer first, particularly in the area of MRT's *clinical validity*. This can be accomplished by answering the question, "How does the use of MRT influence patient outcomes?" In other words, studies are needed of this structure: "Is [MRT technique] effective at influencing the symptoms of [condition] in [patient population] compared to [another intervention]." For example:

1. Is HeartSpeak effective at influencing the severity of panic attacks in panic-prone but otherwise healthy adults, compared to CBT?
2. Is Nutritional Response Testing mediated by a kinesiologist (who uses MRT) more effective than nutritional counseling mediated by a nutritionist (who does not use MRT) at decreasing the severity and duration of symptoms of irritable bowel syndrome in women aged 18 to 65 years.

Certainly, more research into the validity of MRT is required in order to determine how and when it is best employed.

Summary

Detractors of MRT explain its mechanism as being an ideomotor effect, likening it to trickery or hypnotic suggestion and placing it alongside the pendulum, dowsing, Facilitated Communication, automatic writing, the Ouija board, and other unexplained phenomena (Hyman, 1999; Carroll, 2015; Barrett, 2014). Recent research suggests that this is likely incorrect. In order for a phenomenon to be an ideomotor effect, the practitioner (or operator) must be suggestive or evoke an influence on the subject (or patient). The results of this study demonstrate that when comparing *blind* and *not blind* conditions, the practitioner evokes no influence, so it is unlikely that the practitioner is responsible for an ideomotor effect. Likewise, the patient has been shown to produce no significant influence either, so it is also unlikely that the patient is responsible for an ideomotor effect. The limitations of this study are those of any retrospective, observational study in that data were not collected to answer the specific research question of this study. Clearly, more prospective research is required. Another explanation of these results may lie in the colloquial use of inaccurate or inconsistent terminology, which could be resolved by clearer definitions. In the interim, the ideomotor explanation of MRT should be regarded as obsolete until such a time as a more plausible explanation of its mechanism of action is established. For now, when asked the question "How does muscle testing work?" the only accurate response is: "We do not yet know."

References

Barrett, S. (2014). Applied kinesiology: Phony muscle-testing for "allergies" and "nutrient deficiencies." Retrieved November 11, 2017, from https://www.quackwatch.org/01QuackeryRelatedTopics/Tests/ak.html

Benchimol, E. I., Smeeth, L., Guttmann, A., Harron, K., Moher, D., Petersen, I., … RECORD Working Committee. (2015). The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. *PLoS Medicine, 12*(10), e1001885.

Bohannon, R. W. (1997). Internal consistency of manual muscle testing scores. *Perceptual and Motor Skills, 85*(2), 736–738.

Bossuyt, P. M, & Leeflang, M. M. (2008). Developing criteria for including studies. In *Cochrane handbook of systematic reviews of diagnostic test accuracy 4.0* (Chapter 6). London, UK: The Cochrane Collaboration.

Bossuyt, P. M., Reitsma, J. B., Bruns, D. E., Gatsonis, C. A., Glasziou, P. P., Irwig, L. M., … Lijmer, J G. (2003a). The STARD statement for reporting studies of diagnostic accuracy: Explanation and elaboration. *Clinical Chemistry, 49*(1), 7–18.

Bossuyt, P. M., Reitsma, J. B., Bruns, D. E., Gatsonis, C. A., Glasziou, P. P., Irwig, L. M., … de Vet, H. C. (2003b). Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD initiative. *British Medical Journal, 326*(7379), 41–44.

Braid, J. (1855). *The physiology of fascination, and the critics criticised.* Manchester, UK: Grant.

Carpenter, W. B. (1852). *On the influence of suggestion in modifying and directing muscular movement, independently of volition*. London, UK: Royal Institution of Great Britain.

Carroll, R. T. (2012). Applied kinesiology. *The skeptic's dictionary*. Retrieved August 8, 2013, from www.skepdic.com/akinesiology.html

Carroll, R. T. (2015). Ideomotor effect. *The skeptic's dictionary*. Retrieved November 11, 2017, from http://skepdic.com/ideomotor.html

Caruso, W., & Leisman, G. (2000). A force/displacement analysis of muscle testing. *Perceptual and Motor Skills, 91*(2), 683–692.

Caruso, W., & Leisman, G. (2001). The clinical utility of force/displacement analysis of muscle testing in applied kinesiology. *International Journal of Neuroscience, 106*(3–4), 147–157.

Conable, K., Corneal, J., Hambrick, T., Marquina, N., & Zhang, J. (2006). Electromyogram and force patterns in variably timed manual muscle testing of the middle deltoid muscle. *Journal of Manipulative and Physiological Therapeutics, 29*(4), 305–314.

Dillon, K. M., Fenlason, J. E., & Vogel, D. J. (1994). Belief in and use of a questionable technique, Facilitated Communication, for children with autism. *Psychological Reports, 75*(1 Pt. 2), 459–464.

Drouin, J. M., Valovich-McLeod, T. C., Shultz, S. J., Gansneder, B. M., & Perrin, D. H. (2004). Reliability and validity of the Biodex System 3 Pro isokinetic dynamometer velocity, torque and position measurements. *European Journal of Applied Physiology, 91*(1), 22.

Emma, S. (2008). Pitfalls and promises: The use of secondary data analysis in educational research. *British Journal of Educational Studies, 56*(3), 323–339.

Florence, J. M., Pandya, S., King, W. M., Robison, J. D., Baty, J., Miller, J. P., … Signore, L. C. (1992). Intrarater reliability of manual muscle test (Medical Research Council scale) grades in Duchenne's muscular dystrophy. *Physical Therapy, 72*(2), 115–122.

Frost, R., & Goodheart, G. J. (2013). *Applied kinesiology: A training manual and reference book of basic principles and practices* (Rev. ed.). Berkeley, CA: North Atlantic Books.

Gallo, F. P. (2000). *Energy diagnostic and treatment methods*. New York, NY: W. W. Norton.

Garrow, J. S. (1998). Kinesiology and food allergy. *British Medical Journal, 296*(6636), 1573–1574.

Glasziou, P., Irwig, L., & Deeks, J. J. (2008). When should a new test become the current reference standard? *Annals of Internal Medicine, 149*(11), 816–821.

Haas, M., Peterson, D., Hoyer, D., & Ross, G. (1994). Muscle testing response to provocative vertebral challenge and spinal manipulation: A randomized controlled trial of construct validity. *Journal of Manipulative and Physiological Therapeutics, 17*(3), 141–148.

Hall, S., Lewith, G., Brien, S., & Little, P. (2008). A review of the literature in applied and specialised kinesiology. *Complementary Medicine Research, 15*(1), 40–46.

Hyman, R. (1999). The mischief-making of ideomotor action: How people are fooled by ideomotor action. *Scientific Review of Alternative Medicine, 3*(2), 34–43.

Jackson, J. (2005). *The ideomotor effect: A natural explanation for many paranormal experiences*. Retrieved November 11, 2017, from http://www.critical-thinking.org.uk/psychology/the-ideomotor-effect.php

Jacobs, G. E. (1981). Applied kinesiology: An experimental evaluation by double blind methodology. *Journal of Manipulative and Physiological Therapeutics, 4*, 141–145.

Jacobs, G. E., Franks, T. L., & Gilman, P. G. (1984). Diagnosis of thyroid dysfunction: Applied kinesiology compared to clinical observations and laboratory tests. *Journal of Manipulative and Physiological Therapeutics, 7*(2), 99–104.

Jensen, A. M. (2012). Muscle testing (kinesiology): Panacea or placebo? *The Conversation*. Retrieved May 3, 2018, from http://theconversation.com/muscle-testing-kinesiology-panacea-or-placebo-11075

Jensen, A. M. (2015a). *The accuracy and precision of kinesiology-style manual muscle testing*. Oxford, UK: Department of Continuing Education and Department of Primary Health Care Sciences, University of Oxford.

Jensen, A. M. (2015b). Estimating the prevalence of use of kinesiology-style manual muscle testing: A survey of educators. *Advances in Integrative Medicine, 2*(2), 96–102.

Jensen, A. M., Stevens, R. J., & Burls, A. J. (2016). Estimating the accuracy of kinesiology-style manual muscle testing: two randomised-order blinded studies. *BMC Complementary and Alternative Medicine, 16*, 492.

Jepsen, J. R., Laursen, L., Larsen, A., & Hagert, C. G. (2004). Manual strength testing in 14 upper limb muscles: A study of inter-rater reliability. *Acta Orthopaedica Scandinavica, 75*(4), 442–448.

Kenney, J. J., Clemens, R., & Forsythe, K. D. (1988). Applied kinesiology unreliable for assessing nutrient status. *Journal of the American Dietetic Association, 88*(6), 698–704.

King, M. F., & Bruner, G. C. (2000). Social desirability bias: A neglected aspect of validity testing. *Psychology and Marketing, 17*(2), 79–103.

Klinkoski, B., & Leboeuf, C. (1990). A review of the research papers published by the International College of Applied Kinesiology from 1981 to 1987. *Journal of Manipulative and Physiological Therapeutics, 13*(4), 190–194.

Knottnerus, J. A., & Buntinx, F. (2009). *The evidence base of clinical diagnosis: Theory and methods of diagnostic research*. Oxford, UK: Blackwell.

Krebs, C. T., & McGowan, T. O. N. (2013). *Energetic kinesiology: Principles and practice*. Edinburgh, Scotland: Handspring.

Ladeira, C. E., Hess, L. W., Galin, B. M., Fradera, S., & Harkness, M. A. (2005). Validation of an abdominal muscle strength test with dynamometry. *Journal of Strength and Conditioning Research, 19*(4), 925.

Lawson, A., & Calderon, L. (1997). Interexaminer agreement for applied kinesiology manual muscle testing. *Perceptual and Motor Skills, 84*(2), 539–546.

Leboeuf, C., Jenkins, D. J., & Smyth, R. A. (1988). Sacro-Occipital Technique: The so-called arm-fossa test. Intra-examiner agreement and post-treatment changes. *Journal of the Australian Chiropractic Association, 18*, 67–68.

Martin, E. G., & Lovett, R. W. (1915). A method of testing muscular strength in infantile paralysis. *JAMA, 65*(18), 1512–1513.

McGrath, R. E., Mitchell, M., Kim, B. H., & Hough, L. (2010). Evidence for response bias as a source of error variance in applied assessment. *Psychological Bulletin, 136*(3), 450–470.

Moncayo, R., Moncayo, H., Ulmer, H., & Kainz, H. (2004). New diagnostic and therapeutic approach to thyroid-associated orbitopathy based on applied kinesiology and homeopathic therapy. *Journal of Alternative and Complementary Medicine, 10*(4), 643–650.

Monti, D. A., Sinnott, J., Marchese, M., Kunkel, E. J., & Greeson, J. M. (1999). Muscle test comparisons of congruent and incongruent self-referential statements. *Perceptual and Motor Skills, 88*(3), 1019–1028.

Nahmani, L., Serviere, F., & Dubois, J. M. (1984). Kinésiologie de l'articulation temporomandibulaire: Un nouveau test musculaire pour contrôler la normalité de l'occlusion [Kinesiology of the temporomandibular joint: A new muscle test to evaluate the normality of occlusion]. *Cahiers de prothese, 12*(48), 139–148.

Omura, Y. (1981). New simple early diagnostic methods using Omura's "Bi-digital O-ring Dysfunction Localization Method" and acupuncture organ representation points, and their applications to the "drug & food compatibility test" for individual organs and to auricular diagnosis of internal organs—Part I. *Acupuncture and Electro-Therapeutics Research, 6*(4), 239–254.

Ondobaka, S., & Bekkering, H. (2012). Hierarchy of idea-guided action and perception-guided movement. *Frontiers in Psychology, 3*, 579.

Perry, J., Ireland, M. L., Gronley, J., & Hoffer, M. M. (1986). Predictive value of manual muscle testing and gait analysis in normal ankles by dynamic electromyography. *Foot and Ankle, 6*(5), 254–259.

Perry, J., Weiss, W. B., Burnfield, J. M., & Gronley, J. K. (2004). The supine hip extensor manual muscle test: A reliability and validity study. *Archives of Physical Medicine and Rehabilitation, 85*(8), 1345–1350.

Peterson, K. B. (1996). A preliminary inquiry into manual muscle testing response in phobic and control subjects exposed to threatening stimuli. *Journal of Manipulative and Physiological Therapeutics, 19*(5), 310–316.

Pollard, H., Bablis, P., & Bonello, R. (2006). Can the ileocecal valve point predict low back pain using manual muscle testing? *Chiropractic Journal of Australia, 36*(2), 58–62.

Pollard, H., Calder, D., Farrar, L., Ford, M., Melamet, A., & Cuthbert, S. (2011). Inter examiner reliability of manual muscle testing of lower limb muscles without the ideomotor effect. *Chiropractic Journal of Australia, 41*(1), 23–30.

Pollard, H., Lakay, B., Tucker, F., Watson, B., & Bablis, P. (2005). Interexaminer reliability of the deltoid and psoas muscle test. *Journal of Manipulative and Physiological Therapeutics, 28*(1), 52–56.

Pothmann, R., Hoicke, C., Weingarten, H., & Lüdtke, R. (2001). Evaluation of applied kinesiology in nutritional intolerance of childhood. [Evaluation der klinisch Angewandten Kinesiologie bei nahrungsmittel-unverträglichkeiten im kindesalter.]. *Forschende Komplementärmedizin und Klassische Naturheilkunde [Research in Complementary and Natural Classical Medicine], 8*(6), 336–344.

Rolfes, A. E. (1997). *The phenomenon of indicator muscle change*. Newrybar, NSW, Australia: School of Health and Human Sciences, Southern Cross University.

Schmitt, W. H., & Cuthbert, S. C. (2008). Common errors and clinical guidelines for manual muscle testing: "The arm test" and other inaccurate procedures. *Chiropractic and Osteopathy, 16*(1), 16.

Schmitt, W. H., & Leisman, G. (1998). Correlation of applied kinesiology muscle testing findings with serum immunologlobulin levels for food allergies. *International Journal of Neuroscience, 96*(3–4), 237.

Schwartz, S. A., Utts, J., Spottiswoode, S. J. P., Shade, C. W., Tully, L., Morris, W. F., & Nachman, G. (2014). A double-blind, randomized study to assess the validity of applied kinesiology (AK) as a diagnostic tool and as a nonlocal proximity effect. *Explore: The Journal of Science and Healing, 10*(2), 99–108.

Shin, Y. K., Proctor, R. W., & Capaldi, E. J. (2010). A review of contemporary ideomotor theory. *Psychological Bulletin, 136*(6), 943–974.

Slezák, P., & Waczulíková, I. (2011). Reproducibility and repeatability. *Physiological Research, 60*(1), 203–204.

Stock, A., & Stock, C. (2004). A short history of ideo-motor action. *Psychological Research, 68*(2–3), 176–188.

Teuber, S. S., & Porch-Curren, C. (2003). Unproved diagnostic and therapeutic approaches to food allergy and intolerance. *Current Opinion in Allergy and Clinical Immunology, 3*(3), 217–221.

Thie, J., & Thie, M. (2005). *Touch for health: A practical guide to natural health.* Camarillo, CA: DeVorss.

Tiekert, C. G. (1981). Applied kinesiology: Its use in veterinary diagnosis. *Veterinary Medicine, Small Animal Clinician, 76*(11), 1621–1623.

Triano, J. J. (1982). Muscle strength testing as a diagnostic screen for supplemental nutrition therapy: A blind study. *Journal of Manipulative and Physiological Therapeutics, 5*(4), 179–182.

Tripathy, J. P. (2013). Secondary data analysis: Ethical issues and challenges. *Iranian Journal of Public Health, 42*(12), 1478–1479.

Wadsworth, C. T., Krishnan, R., & Sear, M. (1987). Intrarater reliability of manual muscle testing and hand-held dynametric muscle testing. *Physical Therapy, 67*(9), 1342–1347.

Walther, D. S. (1981). *Applied kinesiology, Volume 1: Basic procedures and muscle testing* (2nd ed.). Pueblo, CO: Systems DC.

Walther, D. S. (1988). *Basic procedures and muscle testing in applied kinesiology*, Vol. 1. Pueblo, CO: Systems DC.

Walther, D. S. (2000). *Applied kinesiology: Synopsis* (2nd ed.). Pueblo, CO: Systems DC.

Wingett, M. (2013). Ideomotor response: Unconscious signals and convincers. Retrieved May 3, 2018, from http://www.nlplifetraining.com/general-articles-ideomotor-response-unconscious-signal-convincers