Advances in Integrative Medicine 7 (2020) 29–38

Contents lists available at ScienceDirect

Advances in Integrative Medicine

journal homepage: www.elsevier.com/locate/aimed

Original Research Paper

Investigating the validity of muscle response testing: Blinding the patient using subliminal visual stimuli

Anne M. Jensen[a,*], Richard J. Stevens[b], Amanda J. Burls[c]

[a] Department of Continuing Professional Education and Department of Primary Care Health Sciences, University of Oxford, Oxford, United Kingdom
[b] Department of Primary Care Health Sciences, University of Oxford Oxford, United Kingdom
[c] School of Health Sciences, City University London, London, United Kingdom

ARTICLE INFO

Article history:
Received 4 August 2018
Received in revised form 13 February 2019
Accepted 4 April 2019
Available online 7 June 2019

Keywords:
Kinesiology
Muscle weakness
Lie detection
Deception
Lying
Arm
Upper extremity
Subliminal
Nonconscious
Preconscious
Consciousness

ABSTRACT

Research objective: To determine if Muscle Response Testing (MRT) can be used to distinguish lies from truths using blind test patients.
Design: A prospective study of diagnostic test accuracy was carried out using MRT to distinguish lies from truth.
Methods: Twenty practitioners who routinely practised MRT were paired with 20 blind test patients (TPs). TPs were asked to speak simple true and false statements about visual stimuli presented subliminally (at 20 ms). In the *subliminal phase*, pairs performed 20 MRTs and 20 Intuitive Guesses (IG), consisting of 2 blocks of 10 statements each. In the *Supraliminal Phase*, the same picture-statement pairs were repeated. The order of stimuli presentation was randomly assigned so that each pair was presented with a unique series of stimuli.
Results: In the *Subliminal Phase*, MRT accuracy (as percent correct) was found to be 48.5% (95% CI 42.8–54.2), which was no different from IG accuracy (47.8%; 95% CI 43.2–52.3; $p = 0.68$) or chance (50.0%; $p = 0.59$), and no different from MRT accuracy during the *supraliminal phase* (59.0%; 95% CI 50.4–67.6; $p = 0.05$). However, *supraliminal* MRT accuracy was significantly different from chance ($p = 0.04$), indicating that the pairs could perform MRT proficiently.
Conclusion: The main reason for finding no effect is suspected to be due to an inadequate subliminal methodology, a process which is quite complex. Other explanations of results include: (1) MRT is not a valid test when the TP is blind, (2) Blinding TPs during MRT will produce ambiguous or unpredictable results, or (3) Nonconscious beliefs cannot be elicited using subliminal stimuli. Future research may wish to focus on exploring these possibilities. More specifically, subsequent studies may wish to use different methods to blind TPs, and establish whether MRT can be used to detect nonconscious processes, a generally held consensus among MRT practitioners.

What is already known about the topic

Muscle Response Testing (MRT) is used by over 1 million people worldwide, most commonly within the field of complementary and alternative medicine.

MRT has been shown to accurately detect lies using verbal statements in test patients who were not blind to the verity of the statements they were speaking.

In clinical practice, MRT is routinely used to detect “nonconscious beliefs” and to elicit information about a patient of which the patient is not conscious.

Nonconscious beliefs, for example, in the form of prejudice, do indeed exist.

What this paper adds

Despite its negative results, this paper provides a methodological framework for future studies on MRT.

This paper discusses ways in which nonconscious beliefs may be explored using MRT in the future, taking into consideration its methodological strengths and weaknesses.

These results support the findings of previous studies in this series showing that MRT can be used to accurately distinguish lies from truth using supraliminal stimuli.

1. Introduction

Muscle Response Testing (MRT) is a common assessment method used in complementary and alternative medicine (CAM), and is estimated to be used by over 1 million people worldwide {Jensen, 2015 #4099}. Types of practitioners who may use MRT in a

* Corresponding author at: Wolfson College Oxford, OX2 6UD, United Kingdom.
E-mail address: dranne@drannejensen.com (A.M. Jensen).

https://doi.org/10.1016/j.aimed.2019.04.002
2212-9588/

clinical setting include (but are not limited to): kinesiologists, chiropractors, osteopaths, psychologists, naturopaths and others – however, not all practitioners of these types employ MRT in their practices – only those who have pursued specific training in it.

Many practitioners report that MRT is indeed one of the biggest strengths of their practice – because it can be used to pinpoint the source of problems quickly – and yet it is also one of the biggest weaknesses – due to its lack of scientific validation. Possibly because of this lack of validation, outside of the CAM arena and among those who have little experience with MRT, it is poorly understood and looked upon with abject skepticism – and perhaps rightly so. Despite its widespread use {Jensen, 2015 #4099}, in reality, MRT has poor face validity, and there is little evidence to support its use to accurately detect any condition; as such, there exists a considerable need for rigorous research in this area. This study is one in a series of studies assessing the accuracy and precision of MRT used in a specific way: to detect a false spoken statement (i.e. to distinguish lies from truth) – a target condition which is used consistently among MRT practitioners in many fields.

It differs distinctly from the other two types of manual muscle testing (MMT) utilised in health care today: orthopedic-neurological style of MMT (ON-MMT) and Applied Kinesiology (AK) style of MMT (AK-MMT). In ON-MMT, established by Kendall et al. {Kendall, 1949 #1574;Kendall, 2005 #4199}, a practitioner tests any muscle for the purpose of assessing its strength with the aim of detecting an improvement or a decline in a neuromusculoskeletal condition (e.g. polio, or neuropathy associated with spinal disc degeneration). Its outcome is rated on a 0 to 5 scale (with 5 being normal). In contrast, AK-MMT, developed by Dr George Goodheart {Frost, 2013 #4123}, is a binary test, meaning it only has two possible outcomes, conventionally termed "strong" and "weak." Like with ON-MMT, with AK-MMT the practitioner also may test any muscle, yet the interpretation of the outcome of the test is not limited to the neuromusculoskeletal system; it is dependent upon the muscle being tested. For example, if the popliteus muscle is deemed to be "weak," this may indicate the presence of a gall bladder condition, or alternatively any number of other unconventional conditions, such as an imbalance in the gall bladder meridian {Walther, 1981 #256}.

MRT differs from these other types of MMT, in that it uses only one muscle for testing – often called *the indicator muscle* – as opposed to testing all muscles of the body. Generally, in MRT the importance lies in what the practitioner is aiming to detect rather than the choice of indicator muscle (i.e. theoretically, any muscle can be used as an indicator muscle). Stemming from AK-MMT, MRT does share some of its same characteristics. Namely, MRT is also a binary test and it also tests for conditions beyond the neuromusculoskeletal system.

During an MRT test, a practitioner applies pressure to the indicator muscle until s/he ascertains if the muscle will hold (i.e. tests "strong") or not (i.e. goes "weak") – usually within 1–2 seconds {Thie, 2005 #1577}. The practitioner tests repeatably to detect the presence or absence of target conditions (one MRT test per target), and the target condition can change from one test to the next. Examples of commonly investigated target conditions include (but are not limited to): stress, organ dysfunction, meridian imbalance, toxicity, hypersensitivity, and nutritional need.

A key feature of rigorous studies of diagnostic test accuracy is the blinding of assessors to prior test outcomes {Bossuyt, 2003 #3391;Bossuyt, 2003 #697;Bossuyt, 2003 #3391}. If an assessor (i.e. a tester or a practitioner) is not blinded, this can lead to an information bias, which may result in an overestimation of accuracy {Roever, 2016 #4131}. While much consideration goes into the methods for blinding assessors (in this study as well), little is written about the blinding of patients during the assessment of diagnostic tests. However, it is thought that response bias can be a genuine concern in clinical research {Nichols, 2008 #4132; Furnham, 1986 #4023;Sackett, 1979 #3613}.

Response bias, defined as a tendency of participants in an experiment to consciously or nonconsciously act in a way that they think the experimenter wants them to act, often occurs when participants are aware of the purpose of the study {[1] #4133}. Response bias may not be a potential threat in all studies of diagnostic test accuracy, but would be a concern when assessing those tests in which the patient has the ability to modify his/her response, such as when assessing MRT. There are ways to minimise the risk of response bias, such as not revealing the study aims to the patients being assessed. Another way is to blind the patients to the outcome of the test. These methods of blinding patients were incorporated in this study.

Previous studies in this series of diagnostic test accuracy studies found that MRT can be used to accurately distinguish lies from truth {Jensen, 2016 #4012;Jensen, 2017 #4129;Jensen, 2018 #4134}. In the first study of this series, 48 practitioner–test patient (TP) pairs performed 60 MRTs with an accuracy rate of 65.9% correct, 95% confidence interval (CI) of 62.3–69.5%, compared to an intuitive guessing accuracy of 47.4% correct (95% CI 44.9–50.0; $p < 0.01$). The TPs recruited into this study were all naïve to MRT, meaning no TP had any prior experience with MRT, and they were blind to the study aims and paradigms. The second study in this series was a replication of the first study, and used a mix of naïve and non-naïve TPs. Enrolled into this study were 20 practitioner-TP pairs, including some of the same participants from the first study. They performed 40 MRTs and 40 Intuitive Guesses, and this study produced similar results (mean MRT accuracy 59.4%; 95% CI 54.1–64.7; mean Intuitive Guessing accuracy 51.4%, 95% CI 48.3–54.4; $p < 0.01$). In addition, the second study found no significant difference in MRT accuracy between pairs with a naïve TP and pairs with a non-naïve TP. Notably, these two studies achieved similar, consistent and encouraging results regardless of the naivety of the TP.

When attempting to minimise bias in a clinical investigation, it is especially important to introduce various levels of blinding in the methodology. In both of the studies reported above, the TPs were blind to the study aims and paradigms, and were muscle tested by a practitioner after s/he spoke a given true or false statement. While the practitioner was blind to the verity of the statements, TPs were not: they knew when the statements they were speaking were true and false. They were also not blind to the test outcome: that is, they could observe the outcome of the muscle test (as being *weak* or *strong*). Since TPs were not blind in these two fundamental ways, there was the possibility of them introducing bias during these previous studies. Ideally, to eliminate the likelihood of response bias, the TPs should be fully blinded to test outcomes. This would mean performing MRT after TPs spoke statements in which they did not know were true or false.

Therefore, the objective of this study was to determine if Muscle Response Testing (MRT) can be used to distinguish lies from truths when patients are blinded to the veracity of their statements. It is hypothesised that MRT accuracy when patients are blind will be comparable to when patients are not blind.

2. Methods

This study was a prospective study of diagnostic test accuracy. No participant was assessed prior to enrolment. This protocol received ethics committee approval by the Oxford Tropical Research Ethics Committee (OxTREC; Approval #41-10) and the Parker University Institutional Review Board for Human

Subjects (Approval # R18_10). Also, this study protocol was registered with two clinical trials registries: the Australian New Zealand Clinical Trials Registry (ANZCTR; www.anzctr.org.au), and US-based ClinicalTrials.gov. Written informed consent was obtained from all participants, and all other tenets of the Declaration of Helsinki were upheld. This paper was written in accordance with the Standards for the Reporting of Diagnostic Test Accuracy Studies (STARD) guidelines {[2] #3391 [2]; #697; Bossuyt, 2008 #1525}.

This study followed the same fundamental methodology as the previous studies in this series {Jensen, 2016 #4012;Jensen, 2016 #4012;Jensen, 2018 #4151;Jensen, 2015 #3872}, with a number of modifications for subliminal testing. The primary change was to the stimuli presented to the TP: (1) a different database of pictures & statements was used, and (2) the size of the pictures, the location on the screen and the duration of presentation were modified for this study.

Fig. 1. An example of Muscle Response Testing: A practitioner (right) performs MRT on a patient (left) – using the patient's right deltoid muscle.

2.1. Summary of testing methods

Patients viewed a computer screen on which was displayed a series of pictures. In the first part of the study, the pictures were displayed subliminally, and in the second part, supraliminally. Patients were also given specific instructions to speak either a true statement or a false statement about each picture viewed. The practitioners applied MRT following each spoken statement to determine if it was true or false by using the paradigm that if the MRT outcome was *strong*, this indicated that the statement was *true*, and if the MRT outcome was *weak*, this indicated that the statement was *false*.

2.2. Participants and setting

Healthcare practitioners who routinely use MRT in practice were consecutively recruited as "practitioners" (n = 20). In addition, a mixture of MRT-naïve and MRT-experienced test patients ("TPs") were also consecutively recruited (n = 20 in total). Direct contact (via email or telephone), social media and word of mouth were used to recruit participants, in November 2011, in the US state of California. Volunteers were eligible if they were aged 18–65 years, were fluent in English, and had fully functioning, painfree upper extremities. Volunteers were excluded if they lacked sight, hearing or speech. All practitioners, from any profession, who met the inclusion criteria were enrolled, regardless of the extent or breadth of their MRT experience.

2.3. The primary index test: MRT

During any muscle test, a practitioner applied a force to an extremity which is resisted by a patient using a specific muscle. At first the patient holds the joint in a fixed position, commonly in partial flexion. Then, against the patient's isometric contraction, the practitioner then applies pressure, typically into extension. In this study, practitioners tested the TP's deltoid muscle (see Fig. 1). After performing the MRT, the practitioner, alone, decided if the muscle went "weak" or stayed "strong." The amount of pressure applied often varies from practitioner to practitioner {Schmitt, 2008 #1576}. In addition, the location of the practitioner's testing hand is inconsistent, but is routinely placed on the patient's distal forearm, just proximal to the wrist joint. In this study, practitioners were asked to follow their usual clinical MRT practices.

2.4. The reference standard: actual verity of spoken statement

The reference standard used in this study was the actual truth of the spoken statement, which was always definitively known. Further, it was presumed that all participants inherently knew the difference between True and False statements. Also, the true/false valences of the statements were randomly presented, with approximately half being true and half being false, with each pair being presented with a different sequence.

2.5. The testing scenario & participant flow

TPs viewed pictures on a computer screen which could not be viewed by practitioners. This study was broken up into two phases: (1) the *subliminal phase*, consisting of 2 blocks of 10 statements each of MRT and intuitive guessing (IG), alternating, and (2) the *supraliminal phase*, which consisted of 2 blocks of only MRT of 10 statements each. The subliminal presentation aspect of the design of this study were loosely modelled after a previous study by Miller {[3] #3687}.

Immediately after viewing a picture selected at random by computer and displayed subliminally, the TPs were given instructions via an earpiece, and were inaudible to the practitioners. Therefore, both the practitioners and the TPs were blind to the verity of the statement. No one else was present during the testing.

In addition, this study was broken up into 2 parts: (1) the *subliminal phase*, consisting of 2 blocks of 10 statements each of MRT and intuitive guessing (IG), alternating, and (2) the *supraliminal phase*, which consisted of 2 blocks of only MRT of 10 statements each. All participants were blind to study aims and were not informed of the proportions of true/false statements. Also, all participants completed the pre- and post-testing questionnaires. For the configuration of the testing scenario see Fig. 2, and for the Participant Flow Diagram, see Fig. 3.

2.5.1. Subliminal phase

The stimuli presented in this phase consisted of a picture, an auditory instruction, an auditory attentional prime and a visual attentional prime. The pictures presented were randomly chosen from a database of 40 pictures which were different from those used in previous studies in this series. Also, they were presented for a much shorter amount of time (20 msec), they were presented smaller (no larger than 3 cm x 3 cm), and they were randomly presented around the screen (not in a central position like in previous studies). For examples of the visual stimuli, see Fig. 4. These pictures were paired with simple auditory instructions: "*Say, 'I just saw a* __________.'"

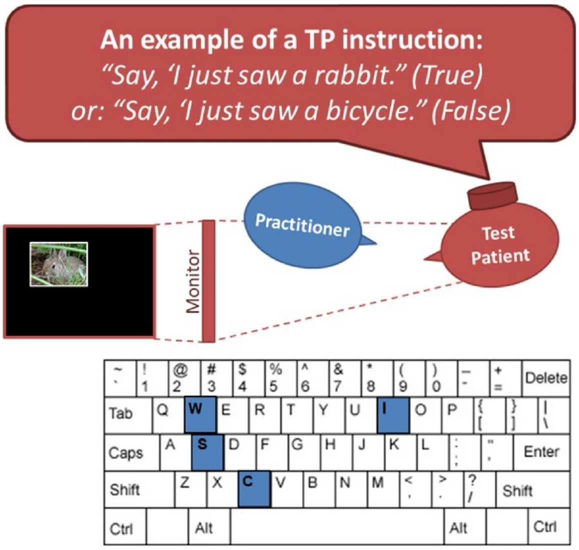


Fig. 2. Testing Scenario Layout: The Test Patient (red) viewed a monitor which the practitioner could see, had an ear piece in his ear through which he received instructions. After the muscle test, the practitioner (blue) entered his results on a keyboard.

The attentional priming stimuli were added to encourage the TP to focus, to ready himself for the presentation of the subliminal stimulus. The auditory attention prime consisted of a 1-second "ding," and was immediately followed by the visual attention prime, an "X" positioned at the center of the screen. The TP was instructed to keep his eyes fixed on the "X," and that the subliminal pictures would be randomly presented around the "X." Following the subliminal picture, the "X" re-appeared. Essentially, the sequence of stimuli presentation was: ["X" – "ding" – Subliminal Picture – "X" – Auditory Instruction]. This was then followed by the TP speaking the statement, the practitioner performing the MRT and then the practitioner entering the result of the MRT, which advanced the TP's screen to a rating scale. See Fig. 5. Since conscious visual perception varies with different stimulus and situational qualities, it must be evaluated on a trial-by-trial basis {[4] #3670}. Unfortunately, this type of rating scale employs subjective report, which, despite being widely used in consciousness trials, may not be ideal in this setting {Irvine, 2012 #3700;Pun, 2012 #3701}. Nevertheless, to appraise conscious perception, TPs were asked to rate how clearly they saw the subliminal picture using a 4-point Likert Scale, anchored with "0 = Didn't see anything" on the left to "3 = Definitely saw it" on the right. Once the TP entered a number from 0 to 3 his screen advanced to the next series of stimuli. This sequence was repeated for 2 alternating blocks each of MRT and IG. Also, for examples of what a portion of the *subliminal phase* might have looked and sounded like to a TP, click here.[a]

The 40 pictures were placed into a database, 20 of which were permanently allocated to the MRT blocks and 20 to the IG blocks. This made it so that all pairs performed MRT after the same 20 pictures, and guessed after the same 20 pictures. However, the order of stimuli was randomly presented using DirectRT™ Research Software (Empirisoft Corporation, New York, NY), so that each pair was presented with a unique sequence of stimuli. Also, the prevalence of lies was again fixed at 0.50, for both the MRT and IG conditions.

2.5.2. Supraliminal phase

In the *supraliminal phase*, pictures were presented like in prior studies in this series. The same 20 pictures presented during the *subliminal* MRT blocks were presented again in a random order during the *supraliminal phase*. The size of the pictures remained consistent, and display location was assigned by the research software to be randomly presented around the screen's center. Once a picture appeared on the screen, it remained until the practitioner completed the MRT and entered his/her result. In this phase the pictures were paired with auditory instructions of this format: "*Say, 'I see a* __________.'" To keep some uniformity between phases a "ding" was also sounded just prior to the auditory instruction. The sequence of stimuli presented during this phase was: "ding" – Supraliminal Picture – 3-second pause – Auditory Instruction. Then the TP spoke the given statement, the practitioner performed the MRT and entered its result, which advanced the TP's screen to next picture-statement pair. In this phase, this sequence was repeated 2 × 10 times, with a short break in the middle (if needed). No IG blocks were included in this phase.

The supraliminal phase was intentionally placed *after* the Subliminal Phase. One reason for this was that the same 20 pictures were used for both phases of MRT (i.e. MRT using *subliminal* pictures and MRT using *supraliminal* pictures), and a stimulus presented first may have had an effect on behavior that follow {[5] #3679;Jaśkowski, 2007 #3631}. For instance, presenting the pictures supraliminally first may have evoked a sort of Mere Exposure Effect during the *subliminal phase*, which may have impacted MRT accuracy {Zajonc, 1968 #3696}. Plus, if the *supraliminal phase* was presented first, TPs might have consciously recognised the pictures during the *subliminal phase*, which would effectively make them not subliminal, thereby negating the point of the study.

2.6. Statistical methods

Based on a previous study in this series in which the accuracy of manual MRT for lie detection had mean 66% and standard deviation 13% across participants {Jensen, 2016 #4012}, we estimated that a sample size of 20 participants would have greater than 99% power to detect an overall accuracy of 66% compared to 50%.

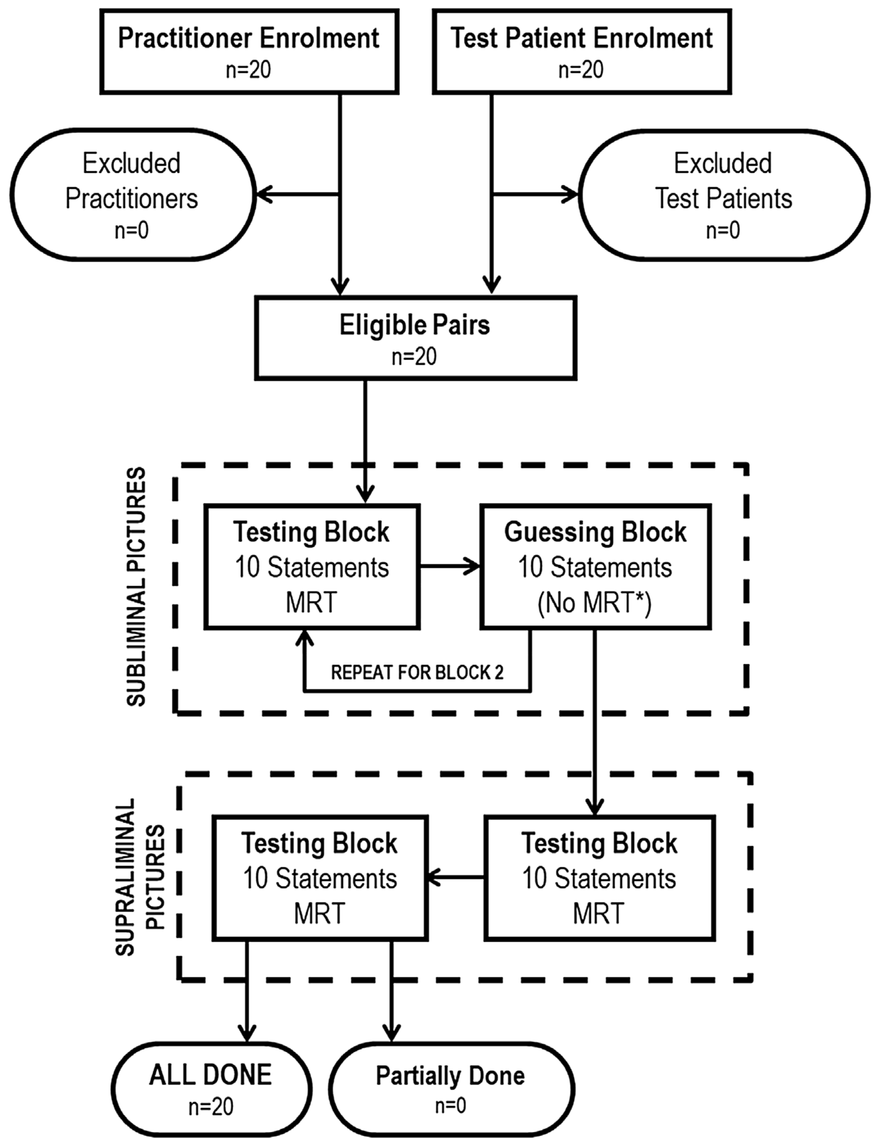
Error-based measures of accuracy will be reported as overall fraction correct {Bossuyt, 2011 #3394} – with the 95% confidence intervals (95% CI). All data were analyzed using STATA 17.0 (StataCorp LP, College Station, Texas), specifically the commands ttest and *pwcorr, sig*.

3. Results

3.1. Participants

Twenty unique practitioner-TP pairs were enrolled, including were 12 female and 8 male practitioners, and 12 female and 8 male TPs. Of the 20 practitioners, there were 16 chiropractors, 2 mental health professionals, and 2 other professionals. Ten practitioners were in full-time practice, and 10 were in part-time practice. The practitioners' mean (SD) number of years in practice was 22.2 (9.4) years. The mean age for practitioners was 53.5 (7.9) years, and for TPs, 38.5 (14.1) years (with 1 T P not responding to this question). For a summary of practitioner demographics, see Table 1 A, and for a summary of Test Patient demographics, see Table 1B (both below).

[a] http://www.drannejensen.com/muscletesting2.html



MRT, Muscle Response Testing; *Touching wrist & observing visual, auditory and kinesthetic cues.

Fig. 3. Participant Flow Diagram.

4. Test results

Pairs took between 20 and 50 min to complete their participation. All pairs completed all testing in full. Aside from TP arm fatigue, there were no adverse events reported from any testing. All accuracies were normally distributed, so parametric statistics were used, mainly the Student t-test.

4.0.1. MRT and IG accuracies

In the *subliminal phase*, the mean (95% CI) accuracy (i.e. overall percent correct) for MRT was 48.5% (42.8–54.2), and the mean (95% CI) IG accuracy was 47.8% (43.2–52.3), which were not found to be statistically different ($p = 0.84$). In addition, the mean MRT accuracy (48.5%; 95% CI 42.8–54.2) was no different from chance (50.0%; $p = 0.59$). See Table 2 below.

The *supraliminal phase* consisted only of MRT (no IG). In this phase, the mean (95% CI) accuracy for MRT was 0.590 (0.504 - 0.676), which was statistically different from chance ($p < 0.01$). When the mean MRT accuracy during the *subliminalphase* (0.485; 95% CI 0.428 - 0.542) is compared to the mean MRT accuracy during the *supraliminal phase (*0.590; 95% CI 0.504 - 0.676), a significant difference was found ($p = 0.04$). See Table 2 (above).

4.0.2. Perceived clarity of perception

Because perception thresholds vary among individuals {[3] #3687}, TPs were also asked to rate how clearly they saw each picture. Even for those trials that TPs reported seeing the picture somewhat clearly (i.e. rating of 2 or 3[b]), MRT accuracy scores were no different from IG ($p = 0.31$) or chance ($p = 0.41$). Comparing this to those trials where TPs reported perceiving little or nothing (i.e. rating of 1 or 0[c]), MRT accuracy scores were still equivalent to IG ($p = 0.94$) or chance ($p = 0.46$).

[b] Ratings: 2 = "I saw the picture and I have somewhat of an idea what it was;" 3 = "I saw the picture and I am sure I knew what it was."

[c] Ratings: 0 = "I saw nothing;" 1 = "I saw something and I have no idea what it was."

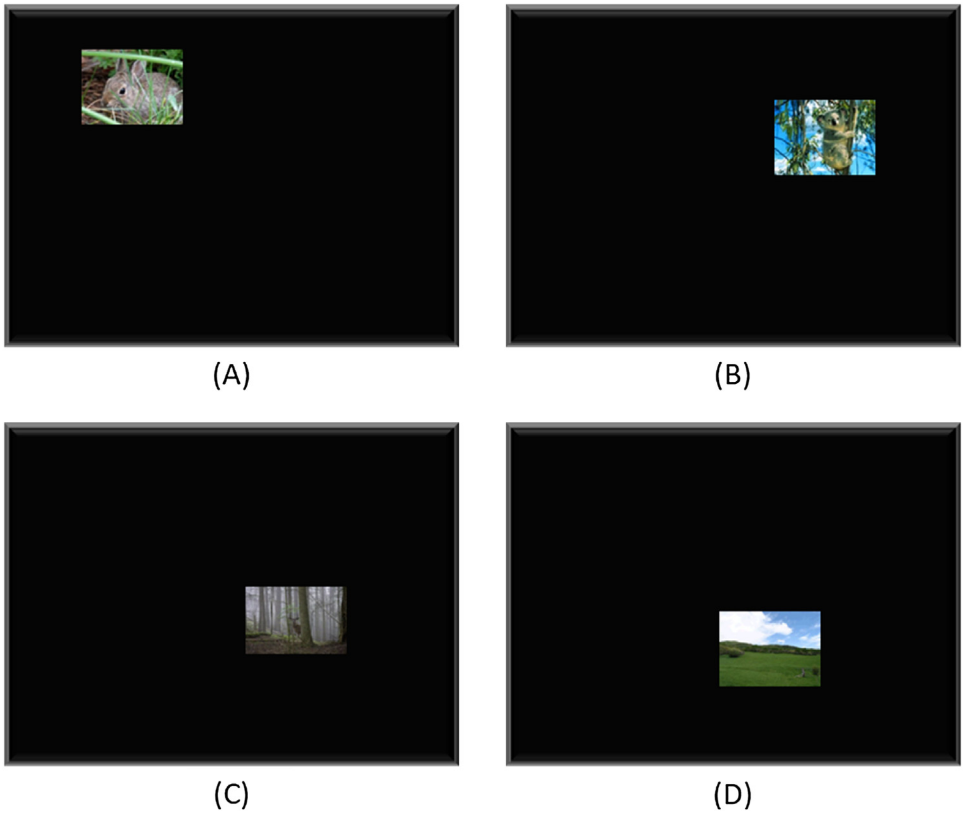


Fig. 4. Examples of Visual Stimuli used during Subliminal Testing. (A), (B), (C) and (D) are examples that could have been presented to a Test Patient during either the MRT or IG Blocks.

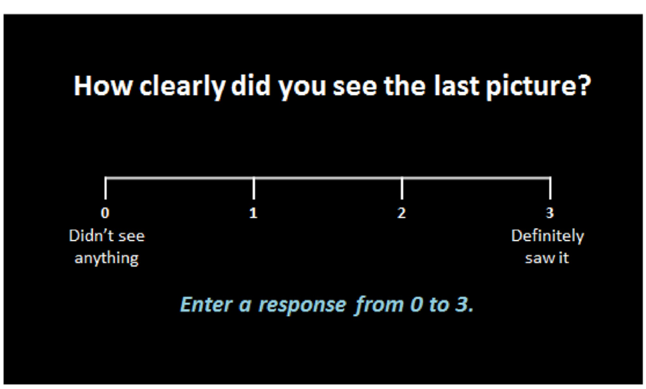


Fig. 5. The Test Patient Picture Rating Scale. Following the MRT, test patients were asked to rate how clearly they saw each subliminal picture flashed on their screen using this rating scale.

5. Discussion

This study failed to demonstrate sufficient MRT accuracy when patients were blinded to the veracity of the statements they were asked to speak. While the fundamental methods kept identical to previous studies in this series, important modifications were made in order to blind TPs. First, a *subliminal phase* was added which consisted of 2 blocks each of MRT and IG using subliminal visual stimuli, which was followed by a *supraliminal phase* in which the same pictures were presented supraliminally and MRT was again performed.

5.1. Statement of principal findings

The primary purpose of this study was to ascertain if MRT could distinguish between true and false statements when TPs were blind to the verity of the statements. The results showed that MRT could not make this distinction with the methods used. No other studies on MRT have included subliminal stimuli, but a number of studies have shown that affective responses can occur outside of conscious awareness {Bernat, 2001 #4174; Bernat, 2001 #2411}.

On the other hand, since the *supraliminal phase* of this study was analogous to the methods of the previous studies in this series, their results can be readily compared: its mean MRT accuracy was significantly different from chance, which supports the findings of previous studies {Jensen, 2015 #3872;Jensen, 2016 #4012;Jensen, 2018 #4151}.

In contrast, our original hypothesis is not supported. The results of this study suggest that MRT cannot be used to distinguish false from true statements when TPs are blind to the verity of the statements they are speaking. To blind the TPs (i.e. to achieve a state where TPs were unsure of the verity of the statements they were asked to speak), pictures were presented to TPs subliminally. Despite the fact that perception thresholds vary among individuals {[3] #3687}, there was no significant difference in MRT accuracies in pairs whose TPs were also asked

Table 1
Demographics. (A) Practitioners; (B) Test Patients.

Table 1A Demographics of Practitioners

	Practitioners
	(n=20)
Gender (M:F)	8:12
Mean age (SD)	53.5 (7.9)
Mean number of years in practice (SD)	22.2 (9.4)
Practitioner-type (*n*)	
Chiropractor	16
Acupuncturist	2
Other	2
Practitioner Practice Status (*n*)	
Full-time	10
Part-time	10
Mean years of MRT experience (SD)	18.9 (8.9)
Mean hours of MRT/day (SD)	6.6 (7.8)
Mean degree of confidence in own MRT ability (pre-testing)[a] (SD)	8.0 (2.2)
Mean degree of confidence in MRT in general (pre-testing)†(SD)	7.4 (2.5)

Table 1B Demographics of Test Patients

	Test Patients
	(n=20)
Gender (M:F)	8:12
Mean age (SD)	38.5 (14.1)
Previous MRT experience (Yes:No)	7:13
Mean degree of confidence in Practitioner (pre-testing)[a] (SD)	7.5 (2.1)
Mean degree of confidence in MRT in general (pre-testing)[a] (SD)	7.7 (1.9)

MRT, Muscle Response Testing; SD, Standard Deviation; Min, Minimum; Max, Maximum; M, Male; F, Female.
[a] Measured using a 10cm Visual Analog Scale, from 0="None" to 10="Most Ever".

to rate how clearly they saw each picture. Even for those trials that TPs reported seeing the picture somewhat clearly (i.e. rating of 2 or 3[d]), MRT accuracy scores were no different from IG or chance. Comparing this to those trials where TPs reported perceiving little or nothing (i.e. rating of 1 or 0[e]), MRT accuracy scores were still indistinguishable from IG accuracy score or chance. This contradicts our original hypothesis that MRT accuracy when patients are blind will be comparable to when patients are not blind. However, the reason for this contradiction remains unclear. In summary, if a subliminal stimulus can produce a muscle response, these methods failed to elicit one that MRT can be used to detect.

5.2. Possible explanations of results

The fact that the results of this study were not what were expected merits reconsideration of the study hypothesis and methodology. During reflection, it seems likely that there are three primary explanations for these unanticipated results. The first is that MRT is not a valid test and cannot be used with any degree of accuracy. The second is that the blinding of patients during a muscle test will produce ambiguous results; in other words, it may be that *unblindedness* of the patient is integral to MRT success. Third, it may be that either nonconscious beliefs themselves do not exist or cannot be aroused using subliminal stimuli. The fourth explanation is that the methodology used for blinding patients in this study was flawed. Each scenario will be discussed.

The first explanation is favoured by MRT detractors many of whom assert that MRT has no semblance of validity whatsoever and is an example of the unsubstantiated dogma which some alternative health movements propagate. While it is agreed that MRT may lack face validity, it is estimated to be practiced by over 1 million people worldwide {Jensen, 2015 #4099}, Despite its prevalent use, research into its validity is in its early stages. Nevertheless, a lack of evidence does not indicate that a test or intervention is not valid, it simply means that there is a lack of evidence and that research is needed. According to Bossuyt, evaluating a new test must be done in three stages, namely assessing its (1) analytical validity, (2) clinical validity, and finally (3) clinical utility {Bossuyt, 2011 #3394}. Since this series of studies represents an attempt at answering the questions, “Is the test true and meaningful?,” these studies are evaluating MRT’s analytical and clinical validity. Because previous studies in this series using a similar methodology showed that MRT could be used to distinguish lies from truth with a significant amount of accuracy, this first explanation seems unlikely {Jensen, 2018 #4134;Jensen, 2016 #4012}. Nevertheless, further research is needed to evaluate MRT’s clinical utility.

The second explanation that blinding test patients produces meaningless results is also unlikely. The reason for this is that in this and in previous studies in this series, similar and adequate accuracies were achieved in pairs whose test patients guessed the paradigm being studied or not (that is, some test patients remained blind and others did not). Future research may wish to further explore the concept of blinding test patients.

The third scenario, that either nonconscious beliefs do not exist or are not aroused by subliminal stimuli, is also unlikely. Research from the field of social psychology has established that nonconscious beliefs do exists, for example, in the form of prejudice or mere exposure bias {Riener, 2017 #4201;Zebrowitz, 2008 #4202}.

It follows, then, that the most plausible explanation for the results of this study is that a flawed methodology was used to blind test patients. It became clear that the way that the subliminal visual stimuli were presented were indeed inadequate. The

[d] Ratings: 2 = “I saw the picture and I have somewhat of an idea what it was;” 3 = “I saw the picture and I am sure I knew what it was.”
[e] Ratings: 0 = “I saw nothing;” 1 = “I saw something and I have no idea what it was.”

Table 2
Comparing accuracies of MRT & Intuitive Guessing (IG). (A) MRT vs. IG during Subliminal Phase; (B) MRT accuracies during Subliminal vs. Supraliminal (C) MRT accuracies of TP reporting guessing the paradigm under investigation vs. not - during Subliminal Phase only; (D) MRT accuracies of MRT-naïve TP vs non-naïve TP – during Subliminal Phase only; and (E) MRT accuracies of TP who knew their paired Practitioner vs. those that did not know their paired Practitioner - during Subliminal Phase only.

				Accuracy		
	Index Test	Comparative Condition	*n*	Mean	95% CI	*p*-value
(A)	MRT	Subliminal Phase	20	48.5	42.8 - 54.2	0.84
	IG	Subliminal Phase	20	47.8	43.2 - 52.3	
(B)	MRT	Subliminal Phase	20	48.5	42.8 - 54.2	**0.04***
	MRT	Supraliminal Phase	20	60.7	48.4 - 73.0	
(C)	MRT	TP reported guessing the paradigm - Subliminal Phase	4	47.5	37.2 - 57.8	0.79
	MRT	TP did not report guessing the paradigm - Subliminal Phase	16	48.8	41.6 - 55.9	
(D)	MRT	TP MRT-naïve - Subliminal Phase	13	50.4	42.0 - 58.8	0.29
	MRT	TP non MRT-naïve - Subliminal Phase	7	45.0	37.4 - 52.6	
(E)	MRT	TP knew Practitioner - Subliminal Phase	3	48.3	29.4 - 67.3	0.97
	MRT	TP did not know Practitioner - Subliminal Phase	17	48.5	41.8 - 55.3	

Accuracy - as percent correct; MRT, Muscle Response Testing; IG, Intuitive Guessing; TP, Test Patient; CI, Confidence Interval.

problem of using visual subliminal stimuli is that the Absolute Visual Threshold is dynamic: it varies with time, with choice of stimuli, with environment and by individual. The numerous factors that can influence the Absolute Threshold are outlined in Table 3. The methods used in this study did not address many of these factors, which would have adversely impacted the study outcomes, and is most likely the cause of the negative results.

Furthermore, results of studies claiming to detect perception of subliminal stimuli appear to be inconsistent, and methods using *visual* subliminal stimuli in particular have a long history fraught with methodological difficulties {Bernat, 2001 #2410; [4] #3670}. One reason for the difficulties may be due to the limited capacity of humans to report visual experiences, which many of these studies rely upon as a measure of perception (i.e. "*Yes, I saw*" or "*No, I did not see*") {[4] #3670}. On the other hand, this also begs the question: Is it possible to be conscious of something, and not able (or willing) to report verbally? Vermeiren and Cleeremans found that when participants lack confidence in their perceptual judgment, they are more likely to fail to report, a condition they call "the underperformance phenomenon" {Vermeiren, 2012 #4200}. However, it is important to keep in mind that one can be conscious of something, and not be able to report or not willing to report. For further discussion about the verbal reporting of visual experiences, see Supplement 1.

In summary, it is suspected that the lack of significant results obtained in this study was mainly because the methodology did not present the visual subliminal stimuli effectively. Without more sophisticated measures, it would be difficult to determine if the stimuli presented in this study reached nonconscious (and accessible) processing.

5.3. Strengths and limitations

This study would have been strengthened by a methodology which controlled all factors listed in Table 3. For instance, presenting the images centrally or randomly but consistently 20° right or left of center might have improved visual perception {[6] #3691}. Also, the Absolute Threshold should have been

Table 3
Some characteristics that may affect Visual Absolute Threshold.

	Stimuli Characteristics	Accounted for in this study:
	Intensity or Brightness	No - varied
	Field Brightness	Yes - held constant
	Size	No - varied slightly, and may have been too small
	Shape	No - varied slightly
	Font & font size	Not Applicable
	Relative clarity	No - varied
	Context	No - varied and may have been too busy
	Interposition	No - varied
	Presence of Emotional Content	No - not tracked
Presentation Characteristics		
	Display time	Yes - held constant
	Display location in visual field	Yes - but varied and possibly detrimental
	Time between stimuli	Yes - held constant
	Use of a prime or mask	No prime or mask was used, but may have been unintentionally masked
Environment Characteristics		
	Lumination	No - varied
	Distance from stimulus to fovea	No - varied
	Monitor-type: CRT vs. LCD	Fixed - LCD (less optimal than CRT)
	Monitor's refresh rate	No - varied
	The presence of distractors	No - varied
	Task instructions	Yes - held constant, but perhaps too demanding to detect stimuli
Participant Characteristics		
	Attention	No - varied, not tracked
	Condition or health status	No - varied, not tracked
	Motivations or expectations	No - not tracked
	Adaptation to the stimulus	No - not tracked
	Previous exposure to stimulus	Yes - held constant; No previous exposure

measured for each TP both prior to testing, and again at the end to confirm a degree of uniformity {[3] #3687}. In addition, the stimuli display times should have been individually tailored for each TP. Another modification that would have strengthened this study could have been the use of a CRT[f] monitor, as opposed to using a laptop's LCD[k], as was the case in this study. This would have stabilised the refresh rate and eliminated the potential for timing errors for which LCDs are renowned (Michael Franklin, personal communication, 22 August 2013).

Furthermore, the study would have been strengthened by showing the pictures multiple times, randomly not showing a picture (e.g. showing a blank screen, similar to how the practitioners were blinded in Study 1) and by using intermittent masking, such as with a checkerboard or a random dot kinetogram[g] (Michael Franklin, personal communication, 24 August 2013). In addition, the pictures chosen for the *subliminal phase* were too complex, and as such, would have required extensive perceptual processing to identify content {Bradley, 2007 #3711}. Other studies have found mixed results regarding how picture complexity affects perception {Bradley, 2007 #3711;de Cesarei, 2011 #3712;Hauswald, 2008 #3714;Shigeto, 2011 #3713}. Nevertheless, perhaps presenting simpler images, such as letters, words or symbols, would have facilitated nonconscious processing. Plus, to ensure the stimuli were not consciously perceived, the addition of a "forced choice test" at the end might have also strengthened this study {Voss, 2008 #3697}. Alternatively, it might have been advantageous to use other types of stimuli, rather than visual, such as auditory or tactile, which seem to have more stable absolute thresholds {[7] #3692;Blankenburg, 2003 #3635; [8] #3689}.

A major strength was its simple yet rigorous design. If the challenges associated with the subliminal presentation of stimuli could be resolved, we maintain that basic methodology which follows the STARD Protocol could be used successfully to assess the validity of other applications of MRT. For example, some practitioners use MRT to detect meridian imbalance. In this instance, the Index Test would again be MRT, the Reference Standard could be pulse diagnosis performed by an experienced Traditional Chinese Medicine practitioner, and accuracy could then be calculated in the same way as the methods of this study.

Another strength of this study was that the data from the *supraliminal phase* confirmed the results of previous studies in this series. These studies have shown that using supraliminal neutral stimuli in a similar set-up, MRT is better than chance at distinguishing truth from lies {Jensen, 2016 #4012}. The present study supports this finding, suggesting that MRT can be successfully investigated using rigorous scientific methods.

Finally, this study may be criticised for its small sample size of 20 practitioner-test patient pairs, despite performing the customary sample size calculation using previous data. Therefore, future researchers may wish to consider these results when performing future sample size calculations.

5.4. Unanswered questions and future research

The primary aim of this study was to investigate if MRT could be used to distinguish lies from truths when TPs were blind to the verity of the statements they were speaking. To blind the TPs we chose to present to them subliminal visual stimuli and then asked them to speak basic true and false statements about the stimuli. Future research may wish to blind the TPs in different ways.

Another important area of research on this topic is to investigate if prior MRT experience and familiarity between patient and practitioner influences accuracy. This study and previous studies in the series found no significant difference in MRT accuracies in pairs whose TPs were MRT-naïve compared to pairs whose TPs were not MRT-naïve, nor in pairs who were acquainted with each other compared to pairs who were not. These results seem to suggest that prior MRT experience (of the TP) and familiarity between testing pairs does not influence accuracy; however, this study may have been underpowered for these subgroup analyses. This is important because in an actual clinical setting, it would be typical for patients to have both prior MRT experience and be acquainted with their practitioner. Therefore, this would be a valuable topic for future research.

In addition, since there's a consensus that MRT is used to detect nonconscious beliefs, future research may want to focus specifically on establishing the validity of this premise. In doing so, investigators may first wish to establish if MRT can detect nonconscious processes, and then if successful, undertake the challenge of determining if MRT can detect conscious beliefs. Finally, if both avenues are successful, an attempt at addressing the primary question is warranted: Can MRT be used to detect nonconscious beliefs?

For a further discussion on the problems that arose during the implementation of this study, see Supplement 2.

6. Conclusion

While this study failed to confirm our hypothesis, it does confirm that the methods used were inadequate or inappropriate for the conditions under investigation. As such, it was a valuable exercise, and will serve to influence future research.

The results failed to show a significant difference between MRT accuracy when the TPs were blind and when they were not blind, between MRT accuracy when the TPs were blind and Guessing accuracy when the TPs were blind, and between MRT accuracy when the TPs were blind and chance. The main reason for finding no effect is likely due to an inadequate methodology for presenting subliminal visual stimuli, a process which is quite complex. Other explanations of results include: (1) MRT is not a valid test when the TP is blind, (2) Blinding TPs during MRT will produce ambiguous or unpredictable results, or (3) Nonconscious beliefs cannot be elicited using subliminal stimuli. Future research may wish to focus on exploring these possibilities. More specifically, subsequent studies may wish to use different methods to blind TPs, and establish whether MRT can be used to detect nonconscious processes, a generally held consensus among MRT practitioners.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Author statement

All authors make substantial contributions to conception and design, and/or acquisition of data, and/or analysis and interpretation of data; all authors participate in drafting the article or revising it critically for important intellectual content; and all authors give final approval of the version to be submitted and any revised version. Concept development: AMJ, AJB; Design: AMJ, RJS, AJB; Supervision: AMJ, RJS, AJB; Data collection: AMJ; Data processing: AMJ, RJS; Analysis/interpretation: AMJ, RJS; Literature search: AMJ; Writing: AMJ; Critical review: AMJ, RJS, AJB; All authors read and approved the final manuscript.

[f] LCD, Liquid Crystal Display; CRT, Cathode Ray Tube.

[g] http://www.drannejensen.com/thelounge/Random_Dot_Kinematogram_(Elliptical).gif

Conflicts of Interest

The authors whose names are listed immediately below certify that they have NO affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

Author names: 1. Richard J. Stevens 2. Amanda J. Burls The authors whose names are listed immediately below report the following details of affiliation or involvement in an organization or entity with a financial or non-financial interest in the subject matter or materials discussed in this manuscript. Please specify the nature of the conflict on a separate sheet of paper if the space below is inadequate.

Author names: 1. Anne M. Jensen – now teaches courses in MRT from which she receives income.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

See attached Conflict of Interest Statement. {[10] #3680 [11]; #3682 [12]; #3673 [13]; #3706 [2]; #3391 [14]; #3677 [15]; #3678 [4]; #3670 [16]; #3674; [17]. #3695 [18]; #3676 [19]; #3686 [5]; #3679 [20]; #566;Mulligan, 2007 #3692 [1]; #3685 [21]; #3707 [22]; #3710 [23]; #3702 [24]; #3703 [25]; #3681 [26]; #3672}

Acknowledgments

The authors would like to thank those who volunteered their time to participate in this study, and to those practitioners who offered the use of their facilities for data collection.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.aimed.2019.04.002.

References

[1] Oxford English Dictionary, Belief Available at: http://oxforddictionaries.com/definition/english/belief (Accessed 22 September 2017), (2017) .
[2] P.M.M. Bossuyt, J.B. Reitsma, D.E. Bruns, et al., The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration, Clin. Chem. 49 (2003) 7–18.
[3] J. Miller, Threshold variability in subliminal perception experiments: fixed threshold estimates reduce power to detect subliminal effects, J. Exp. Psychol. Hum. Percept. Perform. 17 (3) (1991) 841–851.
[4] S. Dehaene, J.P. Changeux, L. Naccache, et al., Conscious, preconscious, and subliminal processing: a testable taxonomy, Trends Cogn. Sci. (Regul. Ed.) 10 (2006) 204–211.
[5] P. Jaśkowski, B. Skalska, R. Verleger, How the self controls its "automatic pilot" when processing subliminal information, J. Cogn. Neurosci. 15 (2003) 911–920.
[6] S. Hecht, S. Schlaer, M.H. Pirenne, Energy, quanta and vision, J. Opt. Soc. Am. 38 (1942) 196–208.
[7] N.W. Mulligan, M. Duke, A.W. Cooper, The effects of divided attention on auditory priming, Mem. Cognit. 35 (2007) 1245–1254.
[8] M. Levine, Fundamentals of Sensation and Perception., 3rd ed., Oxford University Press, London, 2000.
[10] C.N. Alexander, Growth of higher stages of consciousness: Maharishi's vedic psychology of human development, in: C.N. Alexander, E.J. Langer (Eds.), Higher Stages of Human Development. Perspectives on Human Growth., Oxford University Press, Oxford, 1990.
[11] T. Baldwin, in: T. Honderich (Ed.), The Oxford Companion to Philosophy., Oxford University Press, Oxford, 1995, pp. 792.
[12] M.G. Bean, J. Stone, G.B. Moskowitz, et al., Evidence of nonconscious stereotyping of hispanic patients by nursing and medical students, Nurs. Res. 62 (2013) 362–367.
[13] J.E. Bogen, On the neurophysiology of consciousness: 1. An overview, Conscious. Cognit. 4 (1995) 52–62.
[14] N. Dasgupta, S. Asgari, Seeing is believing: exposure to counterstereotypic women leaders and its effect on the malleability of automatic gender stereotyping, J. Exp. Soc. Psychol. 40 (2004) 642–658.
[15] G. Davis, Working Definitions of "Non-Conscious": Commentary on Baars on Contrastive Analysis *Psyche* 1, (1994) . http://www.theassc.org/files/assc/2280.pdf.
[16] S. Doyen, O. Klein, C.L. Pichon, et al., Behavioral priming: It's all in the mind, but whose mind? PLoS One (2012) 7.
[17] S. Freud, An Outline of Psychoanalysis, W.W. Norton, New York, 1949.
[18] J. Glaser, E.D. Knowles, Implicit motivation to control prejudice, J. Exp. Soc. Psychol. 44 (2008) 164–172.
[19] G. Güzeldere, Introduction: the many faces of consciousness: a field guide, in: N. Block, O. Flanagan, G. Güzeldere (Eds.), The Nature of Consciousness: Philosophical Debates., MIT Press, Cambridge, MA, 1997, pp. 1–67.
[20] J.E. LeDoux, The Emotional Brain: the Mysterious Underpinnings of Emotional Life, Touchstone., New York, 1996.
[21] M.I. Posner, Attention: the mechanisms of consciousness, Proc. Natl. Acad. Sci. U.S.A. 91 (1994) 7398–7403.
[22] K.H. Pribram, D. McGuinness, Attention and para-attentional processing: event-related brain potentials as tests of a model, Ann. N. Y. Acad. Sci. 658 (1992) 65–92.
[23] C. Sergent, S. Baillet, S. Dehaene, Timing of the brain events underlying access to consciousness during the attentional blink, Nat. Neurosci. 8 (2005) 1391–1400.
[24] P.U. Tse, S. Martinez-Conde, A.A. Schlegel, et al., Visibility, visual awareness, and visual masking of simple unattended targets are confined to areas in the occipital cortex beyond human V1/V2, Proc. Natl. Acad. Sci. U.S.A. 102 (2005) 17178–17183.
[25] R. Webster, Chapter 11: Exploring the Unconscious: Self-Analysis and Oedipus. Why Freud Was Wrong: Sin, Science and Psychoanalysis, The Orwell Press, London, 2005.
[26] J.N. Williams, Moore-paradoxical belief, conscious belief and the epistemic Ramsey test, Synthese: An International Journal for Epistemology, Methodology and Philosophy of Science 188 (2012) 231–246.