Archives Italiennes de Biologie, 156: 64-86, 2018. DOI 10.12871/00039829201817

Claimed effects, outcome variables and methods of measurement for health claims proposed under regulation (EC) 1924/2006 and related to cognitive function in adults

D. MARTINI[1,A], A. INNOCENTI[2,A], C. COSENTINO[2], G. BEDOGNI[3], I. ZAVARONI[4,5], M. VENTURA[6], D. GALLI[7], P. MIRANDOLA[7], M. VITALE[7], A. DEI CAS[4,5], R.C. BONADONNA[4,5], G. PASSERI[8], C. PRUNETI[2], D. DEL RIO[1]

[1] The Laboratory of Phytochemicals in Physiology, Department of Food and Drugs, University of Parma, Parma, Italy;
[2] Department of Medicine and Surgery, Clinical Psychology Unit, University of Parma, Medical School Building, Parma, Italy;
[3] Clinical Epidemiology Unit, Liver Research Center, Basovizza, Trieste, Italy;
[4] Department of Medicine and Surgery, University of Parma, Division of Endocrinology, Parma, Italy;
[5] Azienda Ospedaliera Universitaria of Parma, Parma, Italy;
[6] Laboratory of Probiogenomics, Department of Chemistry, Life Sciences and Environmental Sustainability, University of Parma, Parma, Italy;
[7] Department of Medicine and Surgery, Sport and Exercise Medicine Centre (SEM), University of Parma, Parma, Italy;
[8] Department of Medicine and Surgery, University of Parma, Building Clinica Medica Generale, Parma, Italy;
[A] Authors equally contributed to the study.

ABSTRACT

Some food/food components have been the object of request of authorization to the use of health claims related to cognitive function in adults and compliant with the Regulation (EC) 1924/2006. Most of the requests have received a negative opinion by the European Food Safety Authority (EFSA) also because of the choice of not appropriate outcome variables (OVs) and methods of measurement (MMs) selected in the trials used to substantiate the claim. This manuscript referes to the collection, collation and critical analysis of OVs and MMs related to cognitive function in adults. OVs and MMs were collected from the EFSA Guidance document and the applications for authorization of health claims pursuant to the Articles 13(5). The critical analysis of OVs and MMs, performed by a literature review, was aimed at defining their appropriateness in the context of a specific claimed effect. The results highlight the importance of an adequate choice of OVs and MMs for an effective substantiation of the claims related to cognitive functioning.
The information provided in this document may serve to EFSA for updating the guidance on the scientific requirements for health claims related to cognitive functions, but also for a better design of randomized controlled trials aimed at substantiating such health claims.

This manuscript has been accepted on December 20, 2017. Part of this manuscript has been also published online at https://efsa.onlinelibrary.wiley.com/doi/abs/10.2903/sp.efsa.2018.EN-1272#

Corresponding authors: Prof. Carlo Pruneti - Tel: +39 0521 904829 - Email: carlo.pruneti@unipr.it; Prof. Daniele Del Rio - Tel: +39 0521 903830 - Email: daniele.delrio@unipr.it

List of abbreviations

ADHD: Attention Deficit Hyperactivity Disorder; **BOLD:** Blood-Oxygen-Level Dependent; **CPT:** Continuous Performance Test; **EEG:** Electro-Encephalo-Gram; **EFSA**: European Food Safety Authority; **ERPs:** Event Related Potentials; **fMRI:** functional Magnetic Resonance Imaging; **GAI:** General Ability Index; **IQ:** Intelligence Quotient; **MCI:** Mild Cognitive Impairment; **MM**: Method of Measurement; **MMSE:** Mini Mental State Examination; **MoCA:** Montreal Cognitive Assessment; **OV**: Outcome Variable; **PET:** Positron Emission Tomography; **PVT:** Psychomotor Vigilance Task; **R 2&7:** Rufus 2 and 7 test; **RAVLT:** Rey Auditory Verbal Learning Test; **RCT:** Randomized Controlled Trial; **RT**: Reaction Time; **RVIP:** Rapid Visual Information Processing; **SCWT**: Stroop Color-Word Test; **WAIS:** Wechsler Adult Intelligence Scale; **WCST:** Wisconsin Card Sorting Test; **WFC:** Word Fragment Completion; **WMS:** Wechsler Memory Scale; **WRAT-4:** Wide Range Achievement Test-4

Introduction

Several evidence demonstrate that cognitive process and food consumption are strictly associated. From one hand, cognitive processes are important determinants of our responses to food cues. An urge to consume a tempting food may be resisted if, for example, a person has a longer term goal of weight loss. There is also evidence that responses to food cues can be facilitated or inhibited by lots of different memory processes. Recent evidence on the influence of working memory and episodic memory processes on responses to food cues was proved (Cammisuli et al. 2017, Higgs 2015, Rogers and Hardman 2015). It is argued that processing of food information in working memory affects how much attention is paid to food cues in the environment and promotes the motivation to seek out food in the absence of direct contact with food cues. It is further discussed that memories of specific recent eating episodes may play an important role in directing food choices and influencing when and how much we eat. However, these memory processes are prone to disruption. When this happens, eating behaviour may become more cue-driven and less flexible. In the modern food environment, disruption of cognitive processing of food reward cues may lead to overconsumption, provoking eating disorders, overweight and obesity.

From the other hand, evidence suggest that cognitive function can be influenced by the consumption of specific food components, foods, or dietary patterns. For instance, it has been shown that dietary patterns including plant-based foods were associated with higher cognitive scores (Pearson et al. 2016), with Mediterranean and DASH diets showing promising results(Smith and Blumenthal 2016).

In this scenario, several food and food components have been the object of applications for authorisation of health claims pursuant to Regulation (EC) 1924/2006 related to psychological functions. Most of these applications have received a negative opinion by the European Food Safety Authority (EFSA), due to an insufficient characterization of the food/food component, the choice of a not appropriate claimed effect and/or to an insufficient substantiation of the claim, including the choice of not appropriate outcome variables (OVs) and/or methods of measurement (MMs). In this scenario, a project has been developed with the aim of improving the quality of applications provided by applicants to EFSA, through an appropriate choice of OVs and MMs (Martini et al. 2017b, Martini et al. 2017c, Martini et al. 2017a). This manuscript referes to the collection, collation and critical analysis of OVs and MMs related to cognitive functioning in adults.

Material and methods: search strategy

The manuscript refers to OVs and MMs collected from the relative Guidance document, from the applications for authorization of health claims under Articles 13.5 and 14 of Regulation (EC) 1924/2006 related to psychological and neurological functions, as well as from comments received during public consultations. The OVs and their MMs were considered only if the food/food constituent(s) was sufficiently characterized and the claimed effect, suitably defined, provided a beneficial physiological effect. Following this decision tree, 4 claimed effects with 16 OVs related to cognitive functioning in adults were evaluated under article 13.5.

Similarly to the methods used in Martini et al. (2017), all the MMs proposed for each OV in the scientific opinions and/or in the Guidance documents were included in the evaluation (Martini et al. 2017c). If no methods were proposed or if the proposed method was considered inappropriate, the critical analysis of the best or the most widely used method was included. Then, a series of keywords were defined for each OV and were used to create individual databases of references on PubMed. The critical evaluation of each OV and MM was performed following a review of the literature deriving from the so obtained databases. Each OV and related MM was ranked in one of the following categories: (i) appropriate alone; (ii) appropriate only in combination with other OVs or MMs; (iii) not appropriate per se; (iv) not appropriate in relation to the specific claimed effect proposed by the applicant(s), (v) not appropriate alone, but useful as supportive evidence for the scientific substantiation of the claimed effect.

Results: critical evaluation of outcome variables and methods of measurement

Function health claims art 13 (5)

Maintenance (i.E. Reduced loss) of cognitive functions

Cognitive functioning

A general definition of cognitive functioning includes the processes of thinking, conceptualization, reasoning, etc. Cognitive function includes many domains like memory, attention (concentration), alertness, learning, intelligence, language, and problem solving (executive function). These entire domains are well-defined and measurable psychological constructs that can be assessed by psychometric and neuropsychological tests.
To evaluate the appropriateness of cognitive functioning as outcome variable of maintenance (i.e. reduced loss) of cognitive functions, the literature deriving from database #1 was critically evaluated (Table 1).
It is intuitive that a general assessment of cognitive functioning is functional in the measurement of cognitive function. If anything, the problem lies in the evaluation of such a complex multidimensional neuropsychological function and the interrelationships between the various cognitive domines.
Cognitive functioning can be assessed by some multidimensional neuropsychological tests, which measure all or many of the above-cited domains. These neuropsychological tests are validated to assess clinical (i.e. dementia, Alzheimer disease, etc.) and subclinical (i.e. Mild Cognitive Impairment, MCI, etc.) conditions of cognitive function loss and are widely used as screening tools (Davis et al. 2015). The total score of some of these neuropsychological tests (e.g. Montreal Cognitive Assessment, MoCA, Mini Mental State Examination, MMSE) can be used to assess cognitive functions. Cognitive functioning can be also assessed by using neuropsychological test batteries (e.g. Reitan, Luria Nebraska, Milan, Erica, etc) that assess diverse domains of cognitive functioning together by using a standardized sequence (Fray and Robbins 1996).
All neuropsychological tests should be validated for the population involved in the study (i.e. language, age range, gender, clinical condition, etc.), and tailored to the specific study design. The possible confounding effect of practice and education needs to be evaluated. In case of test/retest, neuropsychological tests with alternative forms should be used in order to eliminate bias due to the practice. Statistical analysis must be appropriate and conforming to the distribution of the data. A case-by-case evaluation is recommended.
Based on current literature, cognitive functioning appears to be an appropriate outcome variable for the substantiation of health claims in the context of increase or maintenance of cognitive functions.

Montreal cognitive assessment

MoCA is a rapid screening instrument for mild cognitive dysfunction. It is constituted by a 30-point test covering eight cognitive domains: 1) attention and concentration; 2) executive functions; 3) memory; 4) language; 5) visual-constructional skills; 6) conceptual thinking; 7) calculations; and 8) orientation. MoCA is a cognitive screening instrument and not a diagnostic tool.
Total scores >25 are considered to be indicative of normal cognitive functions, whereas scores of 25 or below indicate a possible MCI. A bonus point is given to individuals with less than 12 years of education and there is a basic form to test illiterate individuals or subjects with lower education. MoCa has been validated in many languages (55) and is characterized by a high sensitivity and specificity.

Tab. 1. - Strategies used for retrieving the literature pertinent with outcome variables and methods of measurement related to cognitive functioning.

DB Number	Syntax	Total articles	Narrative reviews	Systematic reviews/metanalyses	Validation studies	Outcome variables/Risk factors
1	((("Neural activity"[Title/Abstract]) OR "Cognition"[Mesh]) OR Brain Function*[Title/Abstract]) AND "English"[lang] AND "Humans"[Mesh]	115429	16661	2445	696	Cognitive functioning
						Neural activity
2	(((attention [MeSH]) OR "Executive Function"[Mesh]) OR "Attention Deficit and Disruptive Behavior Disorders"[Mesh]) AND English [Language] AND Human [MeSH Terms]	80720	6852	1558	522	Selective attention
						Sustained attention
3	("Alertness"[Title/Abstract]) AND "English"[lang] AND "Humans"[Mesh]	3806	536	64	32	Alertness

MoCA has been preferred to other neuropsychological test (i.e. MMSE) because it was formulated as a screening instrument for MCI and not for the clinical assessment of mental disease (i.e. dementia, Alzheimer etc.). However, its applicability must be evaluated case-by-case according to the experimental design and study population (Zhao et al. 2015).
Due to the consideration that MoCa has not alternative forms, it cannot be used in test/retest studies. However, it can be a useful as a screening tool for the selection of subjects to be included in an intervention study.
In conclusion, MoCA is not an appropriate method to measure changes in cognitive functioning in intervention studies.

Neuropsychological test batteries

Neuropsychological test batteries combine a range of tests designed to assess specific cognitive domains in order to provide an overview of cognitive function. There are many validated test batteries (e.g. Cambridge Neuropsychological Test Automated Battery, Halstead-Reitan neuropsychological battery, Luria-Nebraska Neuropsychological Battery, Vienna Test System-Neuro, etc.), each of them designed to assess a set of cognitive domains (e.g. alertness, attention, memory, intelligence, etc.).
Information about validity, reliability, internal consistence, sensitivity and specificity are available for each of them (Purisch 2001, Ross et al. 2013).
Researchers should choose the appropriate battery depending on the experimental design, the study population and the cognitive aspects they would investigate. Based on current evidence, neuropsychological test batteries are appropriate for measuring cognitive functioning in the context of claims on maintenance of cognitive functions, if properly selected on the basis of the study design and population group.

Neural activity

Since the first half of the last century, neural activity is considered strictly related to cognitive functions. Over the past 40 years, techniques applied to measure brain activity significantly improved, allowing the correlation between neural activity of some brain regions and the various domains of cognitive processes. Measures of neural activity in response of specific stimuli or tasks can indicate the role of different brain regions or neural networks in their regulation. Functional measures of neural activity can be divided in direct (i.e. by Event Related Potentials, ERP) and indirect (i.e. by functional Magnetic Resonance Imaging, fMRI).
To evaluate the appropriateness of neural activity as outcome variable of maintenance (i.e. reduced loss) of cognitive functions, the literature deriving from database #1 was critically evaluated (Table 1).
The improvement of techniques to measure neural activity has demonstrated the power of associating brain regions to specific cognitive functions. However, emerging evidence indicates that the so-called "domain-general" areas engage in multiple functions, differing from "domain-specific" areas,

such as the primary visual cortex, which performs a very specific function. Two fundamental features of brain function enable such broad engagement: time and interconnectivity. Brain areas and associated circuits or networks may be engaged in tasks differently over time: some transiently and some consistently. A fundamental understanding of cognition should therefore take into account the dynamic interconnected nature of brain(Braun et al. 2015). For this reason, functional measures of neural activity are not measures of cognitive functions and/or their domains.

In conclusion, neural activity does not appear to be an appropriate outcome variable for the assessment of cognition, and thus for the substantiation of claims on cognitive functions. Measurements of the neural activity of the brain obtained during the performance of a relevant cognitive task can only be used as supportive evidence for the psychometric assessment of cognitive outcomes.

Functional magnetic resonance imaging

fMRI is a recent imaging technique introduced at the end of past century. fMRI is a non-invasive technique that allows localizing the neural activity of the brain with an excellent spatial resolution. Furthermore, it exploits the hemodynamic changes produced by activation of the neurons in order to identify brain areas action. It has a good temporal accuracy, but not enough for the optimal assessing of the temporal dynamics of neural activity, better evaluated by other technique (e.g. ERPs) (Mulert 2013).

This method is based on the MRI signal change, following the hemodynamic and metabolic response in a region in which there is neuronal activation induced by internal or external stimuli. fMRI measures blood oxygenation changes over time (Blood-Oxygen-Level Dependent, BOLD, signal), linked to neuronal activity that is generated in a particular experimental context(Van Horn and Poldrack 2009). It is not a direct measure, but it allows measuring the brain changes caused by variation in neural activity responding to the various sensory, motor or cognitive task.

Based on current evidence, fMRI can be considered a gold standard for spatial localization of the neural activity of the brain (Kim 2011), and thus it is appropriate for measuring neural activity in response to a specific stimulus and/or task.

Event related potentials

ERPs are recorded with surface electrodes placed on the head, likely to as electro-encephalogram (EEG). While the EEG describes the basal electrical brain activity, ERPs consist in a specific variation of the resulting bioelectric signal to the stimulation of a sensorial or motor network to a stimulus or a task. ERPs are made up oscillations of the electrical potential and have a waveform characterized by a series of positive or negative deflections, defined components. Physiologically, ERPs represent the summation of post-synaptic potentials from populations of synchronously active neurons, located primarily in the cortex. Location of ERPs components allows identifying the cortical area activated in response to a particular stimulus based on three parameters:

- Latency: time between the moment of the stimulus application and the time of onset of the component;
- Topography: location on the skull surface is adjustable in which the maximum amplitude of the component;
- Size: size of the deflection of the component relative to baseline.

Therefore, ERP is a direct measure of neural activity: an electrical potential generated by the firing of cortical neurons in response to a specific event (stimulus or task)(Amodio et al. 2014).

ERPs registration has a very good temporal accuracy, allowing the assessment of the dynamic mechanisms of a neural activation

Based on current evidence, ERPs are appropriate for measuring neural activity, especially dynamic mechanisms of the neural response.

Selective attention

Selective attention is the act of focusing on a particular stimulus and simultaneously ignoring irrelevant information from the surrounding environment. This is fundamental in human cognitive activities and in intellectual functions. This occurs on a daily basis because it is impossible to give attention to every stimulus in the environment. Selective attention selects what stimuli are important when events occur. To evaluate the appropriateness of selective attention as outcome variable of maintenance (i.e. reduced loss) of cognitive functions, the literature deriving from database #2 was critically evaluated (Table 1).

Attention is a restricted resource. The ability to perceive the environmental stimuli is limited in terms of both capacity and duration, so attention acts highlighting the details that must be focused, leaving irrelevant information to the side-lines of perception. Selective attention is the process implicated in focusing on a particular stimulus in the environment for a certain period of time, so selective attention allows us to discard unimportant details focusing on what really matters(Eimer 2014).

Selective attention varies depending on the person and its ability to focus or concentrate. It is also affected by distractions in the environment and by other factors such as fatigue, stress, sleep deprivation and motivation in performing the task. Therefore, the assessment of selective attention should be performed after a comprehensive psychological assessment of the subjects examined(Giesbrecht et al. 2014). Selective attention may be a conscious effort, but it can occur subconsciously as well.

On the basis of current literature, selective attention is involved in many cognitive functions and can be considered an important outcome variable in assessing cognitive functions, but "per se" it is not sufficient for the substantiation of health claims in the context of increase or maintenance of cognitive functions.

Conversely, selective attention appears to be an appropriate outcome variable for the assessment of attention (concentration) and therefore for the substantiation of health claims in the context of maintenance of attention (concentration).

Rufus 2 and 7 test

The Rufus 2 and 7 test (R 2&7) is a good example of visual selective search test. It was developed to measure visual attention (both selective and sustained) throughout visual search selection of relevant stimuli while ignoring distracters.

The test consists of a series of 20 trials of a visual search and cancellation tasks: 10 automatic detection trials and 10 controlled search trials.

The subject examined must detect and mark the two target digits: "2" and "7" among alphabetical letters (automatic detection trials) or among other numbers (controlled search trials) that serve as distractors.

Correct hits and errors are counted for each trial and serve as the basis for scoring the test. Speed scores reflect the total number of correctly identified targets (hits). Accuracy scores evaluate the number of targets identified in relation to the number of possible targets.

R 2&7 Test was standardized and normed for use with adolescents and adults, aged 16-70 years(Caban-Holt et al. 2012). Reliability analyses suggest high internal consistency and high split-half reliability for all 2&7 test measures. Convergent and discriminant validity show that test speed and accuracy scores measure sustained attention.

Factorial validity studies confirm that the R 2&7 test measures both sustained attention and selective attention.

Based on current evidence, the R2&7 test is appropriate for assessing selective attention, as well as sustained attention.

Cancellation tests

The cancellation (barrage) tests are the widely used categorical tasks to assess selective attention. Cancellation tasks involve searching and scanning for a specific digit, letter or symbol (geometric or figurative) targets against a background of distracters (e.g Spinnler & Tognoni test, Brickenkamp test, Mesulam Test etc). Distractors are digit, letters or symbols different from the target. Targets and distractors are mixed and can be arranged either randomly or in organized rows and columns. Cancellation tasks that employ a random arrangement of complex symbols are more difficult and subsequently more sensitive than similar tests that are arranged in organized rows and columns. These tests are also used in the assessment of neglect. Cancellation tests are widely used in the neurological assessment of visuospatial function and selective attention (Lowery et al. 2004). Cognitive domains involved in the cancellation task include sustained and selective attention, psychomotor speed, visual searching and motor coordination.

Scoring are given recording the completion time (in seconds), number of correct and incorrect cross-out targets, and searching strategy (randomized or organized searching).

Cancellation tasks showed good reliability, and good test/retest validity. Their sensitivity and specificity increase when more complex tasks are requested (e.g. complex figures in random background). It appears to be an educational effect, with illiterate and low educated subjects showing greater difficulty in tasks performance(Brucki and Nitrini 2008).

Based on current evidence, cancellation tests are appropriate for assessing selective attention in both adults and children. However, because of the numerous variants of these tests commonly used in research, their use should be considered on a case-by-case basis depending on the study population and study design.

Multisensory attention tasks

In everyday life salient sensory information across multiple sensory modalities (e.g. visual, auditory, tactile, etc) are processed simultaneously. That multisensory system allows a rapidly detection of the environmental stimuli.

In multisensory tasks, attention processes are investigated by examining the combined influence of more sensory stimuli. Generally, the association of visual and auditory stimuli or visual, auditory and tactile stimuli is investigated. In the last case, a two–by-two stimuli summation is performed (i.e. visual-acoustic, visual-tactile, acoustic-tactile). In multisensory tests an attention task (e.g. press a button, a pedal, etc) is associated to the co-administration of two different sensory stimuli(Calvert and Thesen 2004). The response is compared with those obtained by the presentation of the single stimuli and of other stimuli in combination.

Multisensory attention tasks measure the interaction of various sensorial pathways in the attention processes. Generally, validated attention tests adapted to the multisensory protocol with appropriate experimental protocols are used. Multisensory attention tasks have been used to support the multisensory attention network teory in both young and old adults(Mahoney et al. 2011).

The test score is affected by various factors, such as fatigue, stress, sleep deprivation and motivation in performing the task. For these reasons, subjects undertaking these tests should undergo a comprehensive psychological assessment.

Based on current evidence, multisensory attention tasks are appropriate for assessing selective attention. However, their validity in the context of a specific study would depend on the complexity of the study design and on their application to a particular context.

Stroop Color-Word Test

The Stroop effect is an interference in the reaction time (RT) of a task. When the name of a colour (e.g., “blue”, “green”, or “red”) is printed in a colour not denoted by the name (e.g., the word “red” printed in blue ink instead of red ink), naming the colour of the word takes longer and is more subjected to errors than when the colour of the ink matches the name of the colour. It requires more attention to recognize a colour than to read a word, so it takes a little longer. The Stroop effect is used in Stroop Color-Word Test (SCWT) to measure many different cognitive functions, including selective attention and its processing speed.

There are different test variants commonly used. They differ in the number of subtasks, type and number of stimulus, times for the task, or scoring procedures, but all versions have at least two numbers of subtasks (Weiss et al. 2007). The written colour name differs from the colour ink it is printed in. In the first trial, the subject must say the written word, while in the second trial he/she must name the ink colour. However, there can be up to four different subtasks, adding in some cases stimuli consisting of groups of letters or dots printed in a given colour with the participant having to say the colour of the ink; or names of colours printed in black ink that have to be read.

In some test variants, the score is the number of items from a subtask read in a given time, in others it is the time that it takes to complete each of the trials. The number of errors and different derived scores are also taken into account in some versions.

SCWT generally has a good reliability and objective validity, and it could be considered a gold standard, but owing to the high number of the test variants commonly used in research and clinical practice and to the absence of a standardization, its application should be evaluated case by case.

Based on current evidence, SCWT is appropriate for assessing selective attention.

Five Digit Test

The Five Digit Test (5-DT) is a good example of selective and alternating (shifting) attention test. It is a numeric four-step test which applies Stroop paradigm.

Subjects tested must read or count Arabic numerals (1-5) or count stars. 5-DT is composed by four subsets: reading, counting, choosing (selective attention), and switching (alternate attention).

Subsets 1 and 2 involve automatic processes such as reading (1, 2, 3, 4 and 5) and counting (quantities from one to five).

Subset 3 (choosing) involves interference control, since an automatic numerical transcoding has to be inhibited for a controlled one (e.g., stimulus = "1,1,1" and response = "three").
Subset 4 (shifting) involves a set-shift from rules of subset 1 to subset 3 and vice-versa depending on an explicit marker.
Executive scores are calculated for inhibition (choosing – reading) and flexibility (shifting – reading). The used measure to evaluate participants' performance is the time spent to complete the tasks for each subset. Faster times indicate a better performance.
The selective and alternating attention indices of the 5-DT show appropriate ecological validity, as well as good predictive validity and responsiveness(Paiva et al. 2016). However, due to the consideration that 5-DT has no alternative forms, it shows a low applicability in test/retest studies(Chiu et al. 2014).
Based on current evidence, 5-DT is a valid test to assess selective actention, but generally it is not appropriate for assessing changes in intervention studies. However, its applicability should be considered on a case-by-case basis.

Sustained attention (vigilance)

Sustained attention or vigilance is the ability to achieve and maintain the focus of cognitive activity on a given stimulus or task and it is a fundamental component of human cognition. In general, attention can be sustained for a limited time, variable in function of psychophysiological conditions, after which there is a decrement of vigilance. It seems that breaks can refresh the ability to maintain attention over time. Sustained attention is strictly correlated to working memory and executive processes, and influences performance of both (McDowd 2007).
To evaluate the appropriateness of sustained attention (vigilance) as outcome variable of maintenance (i.e. reduced loss) of cognitive functions, the literature deriving from database #2 was critically evaluated (Table 1).
Maintaining concentration over time requires good control of attention, so that distracting events cannot capture attention away from its focus. When a task or activity must be focused for a long period of time, sustained attention comes to play. Sustained attention is a basic requirement for information processing and therefore is implied in cognitive function and development. Vigilance can be measured by psychometric tests but the efficiency in sustained attention is affected by various factors as well as fatigue, stress, sleep deprivation and motivation in performing the task. Therefore, tests on sustained attention should be performed after an appropriate psychophysiological assessment of the subjects examined. In addition, the typology and the frequency of target information presentation and the predictability of target location affect the decrement of attention.
On the basis of current literature, sustained attention can be considered an important outcome variable in assessing cognitive functions, but "per se" it is not sufficient for the substantiation of health claims in the context of increase or maintenance of cognitive functions.
Conversely, sustained attention appears to be an appropriate outcome variable for the assessment of attention (concentration) and for the substantiation of health claims in the context of increase or maintenance of attention (concentration).

Continuous Performance Task

There are a variety of Continuous Performance Tasks (CPTs) based on a simple task that primarily measures the ability of subjects to focus attention and to remain vigilant over time (minutes)(Sadeh et al. 2011). A typical CPT task requires a participant to sustain attention over a continuous stream of stimuli (single letters, geometric shapes or digits that are presented serially) and to respond to a pre-specified target.
CPTs are widely used to assess sustained attention, especially in clinical conditions(Shalev et al. 2011). These tests were originally designed to detect deficits in sustained attention among brain-damaged patients and then have been widely used (with some modifications) in studies of psychiatric disorders(Kahn et al. 2012). CPTs may also be used specifically in supporting Attention Deficit Hyperactivity Disorder (ADHD) diagnosis.
Based on current evidence, CPTs are appropriate for assessing sustained attention. However, since they were developed to test patients, their applicability to the healthy adult populations should be evaluated on a case-by-case basis.

Rapid Visual Information Processing

Rapid Visual Information Processing (RVIP) is a test of sustained attention similar to CPTs. It also requires working memory for its successful execution. Single digits appear in a pseudo- random order at a rate of 100 or 200 digits/min on a computer screen. Subjects must detect target sequences of numbers (i.e. 2-4-6 or 3-5-7) and register response with a button press.

RVIP has been used in a number of human psychopharmacological studies. It is sensitive to dysfunction in the parietal and frontal lobe, but it is also a sensitive measure of general mental performance (Coull et al. 1996).

Furthermore, it must be taken into account that differences in scores could be due to differences in the working memory functioning. Therefore, in both case/control and test/retest studies, visual working memory functions should also be assessed.

Based on current evidence, RVIP is appropriate for assessing sustained attention, but visual working memory functions should be assessed as well to allow a meaningful interpretation of the results.

Psychomotor Vigilance Task

Psychomotor Vigilance Task (PVT) is a sustained-attention RT task that measures the speed with which subjects respond to a visual or auditory stimulus, or both (multisensory task). In PVT, the subject presses a button as soon as a light appears or a sound is emitted, the stimulus will turn on randomly every few seconds for 5-10 minutes. The main measurement of this task is not the RT, but the number of times the button is not pressed when the stimulus is on. The purpose of the PVT is to measure sustained attention by counting the number of lapses in attention of the tested subject (Drummond et al. 2005). The measure of RT can be used to assess alertness.

PVT is among the most widely used measures of behavioural alertness and sustained attention for its high sensitivity and ecological validity (Khitrov et al. 2014). A modified version of PVT has been adopted on the International Space Station to provide crewmembers with feedback on neurobehavioral changes in alertness, vigilance and impulsivity, aiding them to objectively identify when their performance capability is decreased.

Based on current evidence, PVT is appropriate for assessing sustained attention and alertness.

Rufus 2 and 7 tesT

See p. 69.

Alertness

Alertness as cognitive domain indicates a state of enhanced arousal, implying a willingness to receive and process information and the preparation to respond. Since the first half of the last century, alertness and arousal refer to the time elapsing between presentation of a stimuli and the response to it. In this contest, the measure of RT throughout a specific test is a relatively standard procedure in the international scientific literature. There is a well-established relationship between RT and alertness and between alertness and the mental activity required for a task. Moreover, electrophysiological correlates of alertness have also been related to RT(Posner 2008).

To evaluate the appropriateness of alertness as outcome variable of maintenance (i.e. reduced loss) of cognitive functions, the literature deriving from database #3 was critically evaluated (Table 1).

The assessment of alertness involves the measurement of the RT after the presentation of a visual (a light that comes on) or auditory (a sound) stimulus. The stimulus can be simple (Simple RT, go/no-go) or complex (Choice RT). In the first case, there is a stimulus and a task to perform (i.e. pressing a button).

A complete assessment of alertness should be made by comparing among them the different RTs (i.e. subtractive method or addictive method). For example, the subtraction of the times achieved in the simplest tasks to those obtained in a more complex task will give an “estimate” of the additional time required by the additional operation (only present in the most difficult task) to be performed.

It was shown that RT changes with a circadian rhythm and is usually longer in the early morning, declines over the course of the day and rises again during the night, peaking in the early morning. Therefore, time of the day in which RT is measured must be taken into account. Alertness and RTs are also affected by mood, sleep quality and duration, stress, fatigue and motivational state, so these characteristics should be assessed in the experimental design.

On the basis of current literature, alertness can be considered an important outcome variable in assessing cognitive functions, but “per se” it is not sufficient for the substantiation of health claims in

the context of increase or maintenance of cognitive functions.
Conversely, alertness can be measured in humans through various methods and it is an appropriate outcome variable for the substantiation of claims on increase or maintenance of alertness.

Simple reaction time tests

Since the first half of the past century, simple RT tests have been used to measure the time of reaction to a single simple stimulus.
Simple RT tests consist of a simple task, like a button to press, to be performed as soon as possible after a visual or auditory stimulus (i.e. a sound, a geometric figure etc.). The same stimulus is repeated in sequence with a casual time interval randomly determined, so that interval between two stimuli cannot be predicted.
Simple RT is a real measure of time elapsing between the presentation of a stimulus and the relative response(Appelle and Oswald 1974).
RT changes with a circadian rhythm and is usually longer in the early morning, declines over the course of the day and rises again during the night peaking in the early morning. Moreover, RTs are also affected by mood, sleep quality and duration, stress, fatigue and motivational state. Therefore, time of the day in which a simple RT test is performed must be taken into account, and subjects participating should undergo a comprehensive psychophysiological assessment.
An overall assessment of RT (i.e. Simple RT tests, Go/No-Go RT tests and Choice RT tests) would be also required.
The complete assessment of RT (i.e. Simple RT tests, Go/No-Go RT tests and Choice RT tests) can be considered the gold standard in the assessment of alertness, but not each of the single tests (i.e. Simple RT in this case).
Based on current evidence, the simple RT tests are appropriate for measuring RT, but "per se" are not sufficient to assess alertness because an overall assessment of RT (including Simple RT tests, Go/No-Go RT tests and Choice RT tests) is needed.

Go/No-Go RT tests

Go/No-Go RT tests consist of a task slightly more complex than Simple RT. A simple task (i.e. press a button) must be performed responding to a stimulus but not when different other stimuli are presented.
Go/No-Go RT tests measure the time elapsing between the presentation of a stimulus, its identification and the relative response(Figueiro et al. 2016). The subtraction of Simple RT from Go/no-Go RT measures the speed of encoding new information and can estimate time required to identify the stimulus.
RT changes with a circadian rhythm and is usually longer in the early morning, declines over the course of the day and rises again during the night, peaking in the early morning. Moreover, RTs are also affected by mood, sleep quality and duration, stress, fatigue and motivational state. Therefore, time of the day in which a Go/No-Go RT test is performed must be taken into account and subjects participating should undergo a comprehensive psychophysiological assessment.
An overall assessment of RT (i.e. Simple RT tests, Go/No-Go RT tests and Choice RT tests) would also be advisable.
The complete assessment of RT (i.e. Simple RT tests, Go/No-Go RT tests and Choice RT tests) can be considered the golden standard in the assessment of alertness, but not each of the single tests (i.e. Go/No-Go RT in this case).
Based on current evidence, Go/No-Go RT tests are appropriate for measuring RT, but "per se" are not sufficient to assess alertness because an overall complete assessment of RT (including Simple RT tests, Go/No-Go RT tests and Choice RT tests) is needed.

Choice reaction time tests

In choice RT tests there are multiple stimuli, and each of them requires a different response. For example, being multiple buttons available, each corresponding to a letter, each time one of the letters appears on a screen the subject need to press the corresponding button. Choice RT tests measure the time elapsing between the presentation of a stimulus, its identification, its discrimination and the relative response. The response time is longer than Simple RT, being mediated by decision-making processes: there is not only the impulse to press a button as soon as the stimulus appears, but also the associated conflict about which key to press. The difference between Choice RT and Go/No Go RT measures the speed of encoding new information, and can estimate the time required to discriminate the stimulus and carry out a choice in the response(Smith et al. 1999).

RT changes with a circadian rhythm and is usually longer in the early morning, decline over the course of the day and rise again during the night peaking in the early morning. Moreover, RTs are also affected by mood, sleep quality and duration, stress, fatigue and motivational state. Therefore, time of the day in which Choice RT Test is performed must be taken into account and subjects participating should undergo a comprehensive psychophysiological assessment.
An overall assessment of RT (i.e. Simple RT tests, Go/No-Go RT tests and Choice RT tests) would be also advisable.
The complete assessment of RT (i.e. Simple RT tests, Go/No-Go RT tests and Choice RT tests) can be considered the gold standard in the measure of RT, but not each of the single tests (i.e. Choice RT in this case).
On the basis of current evidence, Choice RT tests are appropriate for measuring RT, but "per se" are not sufficient in to assess alertness an overall complete assessment of RT (including Simple RT tests, Go/No-Go RT tests and Choice RT tests) is needed.

Psychomotor vigilance task

See p. 72.

Memory

General memory functioning can be defined as the process in which information is encoded, stored, and retrieved. Therefore, memory is a multidimensional construct and can be subdivided in short and long-term memory. Short-term memory is also defined working memory, while long-term memory can in turn be divided into explicit (also called declarative) and implicit (or procedural) memory. Changes in memory can be assessed using memory battery tests (i.e. Wechsler Memory Scale, WMS). Changes in different memory aspects (e.g. working memory, explicit memory and implicit memory) can be assessed throughout specific and valid neurophysiological tests.
To evaluate the appropriateness of memory as outcome variable of maintenance (i.e. reduced loss) of cognitive functions, the literature deriving from database #4 was critically evaluated (Table 1).
It is intuitive that a general assessment of the various aspects of memory is functional in the assessment of memory as cognitive function. The problem lies in the evaluation of such a complex multidimensional construct, which must take into account the interrelationships between the various memory systems which allowinformation storage and recall processes.
In conclusion, memory can be considered an important outcome variable in assessing cognitive functions, but "per se" it is not sufficient for the substantiation of health claims in the context of increase or maintenance of cognitive functions.
Conversely, memory can be used in healthy populations as an outcome variable for the substantiation of health claims in the context of improvement and/or maintenance (i.e. reduced loss) of memory.

Wechsler Memory Scale

Memory as a multidimensional construct can be assessed through standardized test batteries. WMS is a neuropsychological test designed to measure a series of different memory functions. The current version is the fourth edition (WMS-IV) published in 2009 and designed to be used together with the Wechsler Adult Intelligence Scale (WAIS)-IV(Bouman et al. 2016). WMS-IV is made up of seven subtests: Spatial Addition, Symbol Span, Design Memory, General Cognitive Screener, Logical Memory (I & II), Verbal Paired Associates (I & II), and Visual Reproduction (I & II). A person's performance is reported as five Index Scores: Auditory Memory, Visual Memory, Visual Working Memory, Immediate Memory, and Delayed Memory(Elwood 1991).
Validity of all the index scores has been confirmed by factor analysis(Bouman et al. 2015).
WMS is one of the most widely used instruments to assess memory functioning. WMS-R and WMS-IV are validated for ages 16 to 90 years both in normally developing and aging individuals, as well as in patients with a variety of clinical diagnoses. Both versions show a good internal reliability and ecological validity.
Based on current evidence, the WMS, and particularly the latest two versions (VMS-R and WMS-IV), is the best choice for measuring the general memory functioning, and many of its subscales are appropriate for measuring various memory aspects.

Problem solving

Problem solving is a mental process that, using generic or specific methods, finds solutions to changes in the environment or other kind of problems. Problem solving is considered the most complex of all cognitive functions, and has been defined as a higher-order cognitive process that requires the modulation and control of more routine or fundamental skills(Funke 2010). Problem solving is one of the aspects of the executive function. Executive function is a multifaceted neuropsychological construct consisting of a set of higher-order neurocognitive processes that allow mammals to make choices and to engage in purposeful, goal-directed, and future-oriented behaviour.

To evaluate the appropriateness of problem solving as outcome variable of maintenance (i.e. reduced loss) of cognitive functions, the literature deriving from database #5 was critically evaluated (Table 1).

In current theories of problem solving, a problem is conceptualized as composed of a given state, a desired goal state, and obstacles between the given and the goal state. The nature of human problem solving and its processes have been widely studied by psychology, neuropsychology and cognitive sciences.

Problem solving can be assessed by laboratory-based tasks (i.e. WCST, Tower of Hanoi etc.), but attention, emotions, motivation, memory and psychophysiological state interact with the achievement of the goal. Therefore, before the assessment of problem solving, a complete psychophysiological assessment of the subjects should be performed.

Problem solving can be considered an important outcome variable in assessing cognitive functions and their development, but "per se" it is not sufficient for the substantiation of health claims in the context of increase or maintenance of cognitive functions or cognitive development.

Wisconsin Card Sorting Test

Wisconsin Card Sorting Test (WCST) is a neuropsychological tool that assesses abstract reasoning skills, change of cognitive strategies and problem solving in subjects between 6 and 70 years. In addition to the objective scores relative to the overall success, it also provides specific information on the difficulties encountered in the various tasks.

The WCST consists of 4 stimulus-cards and 128 response-cards (2 decks of 64 cards). The stimulus-cards show a red triangle, two green stars, three yellow crosses and four blue circles. The response-cards show variables differing for:

- number (1 to 4 for paper);
- forms (circles, triangles, crosses or stars);
- color (red, blue, yellow and green).

A deck of 64 response-cards is delivered to the subject, who must match each of them to each stimulus-card, following the criterion that he/she believes right.

Each response-card can be combined with a stimulus-card by only one parameter or a combination of the three parameters. Through feedback from the examiner about the correctness of the response, the subject must discover the criterion for a correct classification.

During the test, the classification criteria will be changed without warning and demanding to develop a new classification strategy(Nyhus and Barcelo 2009).

WCST is considered the gold standard for the measurement of problem solving skills in adults. WCST is also an appropriate method to assess problem solving skills in children from 6 years of age onwards(Chelune and Baer 1986).

Intelligence (IQ)

Intelligence is one domain of cognitive function and can be defined as the ability to learn from experience and to adapt to surrounding environments. Intelligence can be understood in part in terms of brain biology, especially with regard to the functioning of the prefrontal cortex, and correlates with brain size. It seems to be influenced by both genetic and environmental factors. Moreover, heritability varies as a function of socioeconomic status and other environmental factors.

To evaluate the appropriateness of intelligence as outcome variable of maintenance (i.e. reduced loss) of cognitive functions, the literature deriving from database #6 was critically evaluated (Table 1).

Many and diverse theories of intelligence have been formulated. All new theories, however, agree on a multidimensionality of intelligence. Psychometric theories conceptualize intelligence as a sort of "map" of the mind and are based upon analyses of individual differences in subjects' performance on psychometric tests. These theories have been the basis for most conventional tests of intelligence ("IQ tests"). Intelligence Quotient (IQ) is a total score derived from one of several standardized tests designed to assess human intelligence. IQ scores are used for educational placement, to evaluate job applicants and in the assessment of both intellectual disability and the effects after brain injury. The many different kinds of IQ tests include a wide variety of items. Some test items are visual, while many are verbal. Some are based on abstract-reasoning problems, while others concentrate on arithmetic, vocabulary, or general knowledge.

Attention, emotions, motivation, memory, learning and psychophysiological state interact with intelligence(Sternberg 2012). Therefore, before the assessment of intelligence, a complete psychophysiological assessment of the subjects should be performed.

In conclusion, intelligence can be considered an important outcome variable in assessing cognitive functions and cognitive development, but “per se” it is not sufficient for the substantiation of health claims in the context of increase or maintenance of cognitive functions or cognitive development.

Wechsler Adult Intelligence Scale - IV

The WAIS is the most widely used IQ test for both adults and older adolescents (16-90 years old) in the world. The original WAIS (Form I) was published in 1955 by David Wechsler, but it is currently in its fourth edition (WAIS-IV) released in 2008. WAIS was designed on Wechsler’s definition of intelligence: “... the global capacity of a person to act purposefully, to think rationally, and to deal effectively with his environment”.

In this context, intelligence was made up of specific elements that could be isolated, defined, and subsequently measured. However, these individual elements are all interrelated, they aren’t independent with other. In other words, general intelligence is composed by several specific and interrelated functions that can be individually measured to obtain a general picture.

The current version of the test, the WAIS-IV, was standardized in 2008, and is composed of 10 core tests and 5 supplemental tests, with the 10 core tests comprising the Full Scale IQ. There are four index scores representing major components of intelligence:

- Verbal Comprehension Index (VCI)
- Perceptual Reasoning Index (PRI)
- Working Memory Index (WMI)
- Processing Speed Index (PSI)

Moreover, two broad scores are also generated, which can be used to summarize general intellectual abilities:

Full Scale IQ (FSIQ), based on the total combined performance of the VCI, PRI, WMI, and PSI;

General Ability Index (GAI) based only on the six subtests that the VCI and PRI comprise.

With the new WAIS-IV, the verbal/performance subscales from previous versions were removed and replaced by the index scores. The GAI was included, which consists of the Similarities, Vocabulary and Information subtests from the Verbal Comprehension Index and the Block Design, Matrix Reasoning and Visual Puzzles subtests from the Perceptual Reasoning Index. The GAI is clinically useful because it can be used as a measure of cognitive abilities that are less vulnerable to impairments of processing and working memory.

On the basis of current evidence, WAIS-IV can be considered the gold standard among IQ tests(Hartman 2009). Therefore, WAIS-IV is the best choice to measure intelligence in adults(McFarland 2013).

Learning

Learning can be defined as a relatively permanent, long-term change in behaviour resulting by experience or training. It is a complex cognitive process in which many aspects of cognitive function are involved. Human learning may occur as part of education, personal development, schooling, or training. It may be goal-oriented and may be aided by motivation.

To evaluate the appropriateness of learning as outcome variable of maintenance (i.e. reduced loss) of cognitive functions, the literature deriving from database #7 was critically evaluated (Table 1).

Individuals constantly interact with the environment and are influenced by it. This interaction leads to changes or modifications of their behaviour in order to deal effectively with the environment. Human skills, knowledge, habits, interests and personality characteristics are all the result of learning. Learning is therefore a multidimensional process involving a wide range of physical and mental activities, and cannot be explained within a limited framework. Learning in adults is usually assessed by measuring various outcomes, such as reading skills, comprehension, spelling, and problem solving(Bullinaria 1997).

In conclusion, learning can be considered an important outcome variable in assessing cognitive functions, but "per se" it is not sufficient for the substantiation of health claims in the context of increase or maintenance of cognitive functions.

Wide Range Achievement Test - 4

The Wide Range Achievement Test 4 (WRAT4) is an achievement test which measures an individual's ability to read words, comprehend sentences, spell, and compute solutions to math problems. Currently, the test is in its fourth revision. The test was developed in 1941 by the psychologists Sidney W. Bijou and Joseph Jastak. The test is appropriate for individuals aged 5-94 years. The WRAT4 provides two equivalent forms (Blue and Green), which enables retesting within short periods without potential practice effects that occur from repeating the same items. The alternate forms may be also administered together in a single examination. The various editions of the Wide Range Achievement Test (WRAT) have enjoyed widespread use in a variety of settings as a measure of the basic academic skills necessary for effective learning, communication, and thinking: reading and spelling words and performing basic mathematical calculations.

WRAT is a multidimensional achievement test that measures ability in reading words, comprehending sentences, spelling, and computing solutions to math problems. The test is validated in subjects from 5 to 94 years old and is currently in its fourth revision (WRAT-4). The WRAT-4 provides two equivalent forms (Blue and Green), which enables retesting within short periods without potential practice effects, so it is also applicable in the test-retest studies.

The WRAT4 includes four sub-tests:

- Word Reading measures letter and word decoding through identification and word recognition;
- Sentence Comprehension measures ability to gain meaning from words and to comprehend ideas and information contained in sentences using a modified closed technique;
- Spelling measures ability to encode sounds into written form using a dictated spelling format containing both letters and words;
- Maths Computation measures ability to perform basic mathematical computations through counting, identifying numbers, solving simple oral problems and calculating written maths problems.

In addition, WRAT4 also yields a Reading Composite score obtained by combining the Word Reading and Sentence Comprehension scores.

Derived scores were developed for both age and grade referenced groups(Arciuli and Simpson 2012). Standard scores, percentiles, stanines, normal curve equivalents and Rasch scaled scores are provided.

For its multidimensional structure WRAT4 is adeguate to assess not only learning but also language skills, reading, reading comprension and spelling.

WRAT-4 is frequently used for between-subject comparisons regarding learning ability and/or learning disability(Leverett et al. 2002). It is applicable to healthy adult populations and can be considered the gold standard to measure learning and language but as all psychometric tests, it must be validated for the studied population.

Basing on current evidence, WRAT-4 can be considered the gold standard for measuring learning and language. Moreover, the WRAT-4 word reading subtest is the best choice to measure reading word, WRAT-4 Sentence Comprehension subtest is the best choice to measure Reading Comprehension and WRAT-4 subtest Spelling is the best choice to measure Spelling.

Language

Language can be defined as a formal system of communication that involves the combination of words and/or symbols, whether written or spoken, as well as some rules that govern them. One of the more studied aspects of language is reading. Reading is a complex cognitive process in which the subject decodes some symbols (reading words) with the final goal of deriving a meaning (reading comprehension). In this context, reading words is the capability of recognising real words, and reading comprehension is a cognitive multifaceted process in which the subject derives a meaning from the codification of several symbols (reading words). Reading comprehension depends on the ability to recognize words, on cognitive development, on memory and on the educational level. The fMRI has been used to determine the specific neural pathways of activation during the reading comprehension processes.

To evaluate the appropriateness of learning as outcome variable of maintenance (i.e. reduced loss)

of cognitive functions, the literature deriving from database #7 was critically evaluated (Table 1).
Sample reading words tasks include measures of irregular and regular words, experimental words, and real-words identification, whereas, reading comprehension is focused on measures of comprehension of a sentence or a text. Skills in both read words and reading comprehension are used to assess learning and/or language skills and/or disabilities in children and adults. The majority of dependent measures in these domains included reading comprehension and general reading measures(Swanson 2012). Reading skills are clearly related to the educational level of the subject tested, therefore an educational assessment of participants is needed. Reading words can be easily measured by psychometric tests (i.e. WRAT-4, WJPB Reading Cluster etc.), and can be used to measure language in healthy adults(Yarkoni et al. 2008).
Language can be considered an important outcome variable in assessing cognitive functions, but "per se" it is not sufficient for the substantiation of health claims in the context of increase or maintenance of cognitive functions.

Wide Range Achievement Test - 4

See p. 77.

Increasing and/or maintenance (I.E. reduction loss) of attention (concentration)

Selective attention

See p. 68.

Rufus 2 and 7 test

See p. 69.

Categorical search attention tasks

See Cancellation tests, p. 69.

Multisensory attention tasks

See p. 70.

Stroop Color-Word Test

See p. 70.

Five Digit Test

See p. 70.

Sustained attention (vigilance)

See p. 71.

Continuous Performance Test

See Continuous Performance Task, p. 71.

Rapid Visual Information Processing

See p. 72.

Psychomotor Vigilance Task

See p. 72.

Rufus 2 and 7 test

See p. 69.

Increasing and/or maintenance (I.E. reduction loss) of alertness

Alertness

See p. 72.

Simple reaction time tests

See p. 73.

Go/No-Go reaction time tests

See Go/No-Go RT tests, p. 73.

Choice reaction time tests

See p. 73.

Psychomotor Vigilance Task

See p. 72.

Improvement, maintenance or reduction loss of memory

Memory

See p. 74.

Wechsler Memory Scale

See p. 74.

Working memory

Working memory maintains information in an easily accessible state over brief periods (several seconds to minutes). Working memory can be defined as a short-term memory system that allows not only to temporarily hold the information, but also to operate on it, manipulating it, updating it constantly, and putting it at the service of other mental operations. Working-memory capacity is limited and may only hold a small amount of information. This implies that when the system is engaged in a task that absorbs a specific amount of these limited resources, an additional task competitor would have deleterious effects on the execution of both tasks. Working memory consists of a multicomponent system that is responsible for the active maintenance and manipulation of information. The working memory model is composed of two processors, a phonologic loop (verbal working memory) and a visual spatial stretchold (visual-spatial working memory), and of an executive center that processes information. The phonologic loop is the working memory subsystem that holds sequences of digits or syllables for immediate recall. Verbal working memory comprises both a verbal and an acoustic component. The visual-spatial stretchold is a subsystem that is assumed to hold visual and spatial information for short periods, so that it can be used during other cognitive processes such as thinking, remembering or processing tasks. Therefore, visual-spatial working memory is implied in remembering the visual features of an object, such as form and colour ("what"), and the spatial location of the object ("where")(Eriksson et al. 2015).

To evaluate the appropriateness of working memory as outcome variable of improvement, maintenance or reduced loss of memory, the literature deriving from database #4 was critically evaluated (Table 1).

Verbal working memory involves both a "mind's ear" that hears the digits when you read them and a "mind's voice" that repeats them in rehearsal. When verbal information visually presented is encoded, the information is translated into a sound-based ("auditory-phonological") code. This code is something like an internal echo that resounds briefly before declining. To prevent complete decay, an active process must refresh the information. Once the verbal information is spoken internally by the "mind's voice" in rehearsal, it can be again heard by the "mind's ear" and maintained in a phonological store. In this way, a continuous loop plays for as long as the verbal material needs to be maintained in working memory. The first step of the process, translation into a phonological code, is necessary only for visually presented material; access of auditory information, such as speech, to the phonological store is automatic. The model of phonologic loop has been experimentally tested and confirmed.

Visual-spatial working memory seems to create and maintain a representation that persists across the irregular pattern of eye movements characterizing our scanning of the visual world. Moreover, this component creates and maintains visual images of the world, used in recalling, imaging and orientation. It has been shown that spatial tasks can interfere with spatial skills, while a more purely visual activity, such as seeing a sequence of pictures or colour patches, may interfere with the capacity to remember objects or shapes. The presence of similarities between storage of serial order in visual and verbal memory suggests an analogous process, though not necessarily within a single system. Visual working memory is strictly related to visual sustained attention. In fact, visual working memory depends on actively sustained maintenance of relevant sensory representations.

Both the verbal and the visual-spatial components of working memory can be measured by neuropsychological tests (i.e. Digit Span and Corsi Block-Tapping Task respectively).

Attention, emotions, motivation and psychophysiological state interact with working memory processes. Therefore, accurate information about alertness, attention, motivation, and the emotional and psychophysiological condition of the subject should be collected and evaluated before the assessment of working memory.

In conclusion, working memory can be considered an important outcome variable in assessing memory, but "per se" it is not sufficient for the substantiation of health claims in the context of increase or maintenance of memory.

Conversely, working memory appears to be an appropriate outcome variable for the substantiation of health claims in the context of increase or maintenance of working memory in adults and children, but both verbal and visual-spatial subsystems should be evaluated.

Digit span

In a typical test of memory span, a list of random numbers or letters is read aloud or presented on a computer screen at the rate of one per second. At the end, a sequence of numbers or letters must be repeated in the same order in which they were submitted.

The digit span test begins with two to three numbers, increasing by one digit at a time until the subject fails. Recognizable patterns (for example 2, 4, 6, 8) should be avoided.

This kind of test with different digits is used in all WAIS and WMS versions, and in several other neuropsychological standardized test battery and it is extensively used in research and diagnostics for the assessment of verbal working memory(Eriksson et al. 2015).

The assessment of Digit Span requests two different tests: 1. Digits Forward; 2. Digits Backward.

While the participant is asked to enter the digits in the given order in the forward digit-span task, in the backward digit-span task the participant needs to reverse the order of the numbers.

The length of the longest list a person can remember is that person's digit span. The average digit span for normal adults without error is seven plus or minus two(Bowden et al. 2013).

Based on current evidence, Digit Span is considered the gold standard for measuring verbal working memory in adults and children. For the substantiation of health claims in the context of improved working memory, visual spatial working memory (i.e. by the Corsi Block tapping test) should also be assessed.

Corsi block-tapping task

The Corsi block-tapping task requires reproduction of a sequence of movements by tapping blocks as demonstrated by an examiner. The test is traditionally administered using nine cubes positioned on a wooden board, but numerous digital versions have been developed. Participants must reproduce sequences of blocks in the order touched by the examiner, or in backward. If a certain proportion of the sequences is reproduced correctly (usually 1/2, 2/3, or 3/5 of the trials per sequence length), the sequence length increases by one item. The procedure ends when the number of wrong reproductions exceeds the proportion of admissible errors per length. The examiner records either the maximum number of blocks correctly reported or the total number of correct lists(Berch et al. 1998).

Corsi block tapping task has been described as the single most important nonverbal task in neuropsychological research(Brunetti et al. 2014). It is extensively used in research and diagnostics for the assessment of visual-spatial working memory, but it also used to assess spatial attention, as part of test batteries and to support hypotheses about the localization of focal brain lesions(Chechlacz et al. 2014).

Basing on current evidence, Corsi block-tapping task is considered the gold standard for measuring visual-spatial working memory in both adults and children. For the substantiation of health claims in the context of improved working memory, verbal working memory (i.e. by the Digit Span) should also be assessed.

Explicit memory

Explicit memory (also called "declarative memory") is one of the two major components of long-term memory. Explicit memory requires conscious thought, such as naming animals that live in the forest. It is what most people have in mind when they think of "memory," and whether theirs is good or bad. Explicit memory is often associative because the brain works as a unique integrated system, linking memories together. Therefore, explicit memory is a memory that can be intentionally and consciously recalled. Explicit (or declarative) memory provides a way to represent the external world. The explicit memory model has two subsystems: episodic memory and semantic memory.

Episodic memory is autobiographical: it provides a crucial record of own personal experiences, and consists of the ability to re-experience a time- and place-specific event in its original context. Any past event in which people played a part, and which they remember as an "episode" (a scene of events) is episodic. This form of memory appears to be centred in the brain's hippocampus and is strictly connected with the cerebral cortex.

Semantic memory is the component of explicit memory that accounts for general knowledge about the world. Semantic memory is the most stable type of memory, and unlike episodic memory, it is better sustained over time. It seems to be located in the hippocampus and related areas, but some authors think it is widely spread throughout the brain. Semantic memory is the component of long-term memory that can be linked to knowledge.

To evaluate the appropriateness of explicit memory as outcome variable of improvement, maintenance or reduced loss of memory, the literature deriving from database #4 was critically evaluated (Table 1).

Episodic memory is the component of memory that can be linked to the remembrance. How well we record an episodic memory depends on several factors. For example, things that occur in emotionally charged conditions often lead to stronger memories. Another important factor is the strength with which the brain recorded data when the first experience took place. When the brain is able to process what the subject sees, hears, smells, tastes, and feels very quickly and accurately, the memory is recorded with more power, making it easier to recall later.

Semantic memory refers to the capacity to actualise crystallised knowledge (facts, ideas and concepts) stored using semantic organisation (semantic network) and conceptualisation processes. All semantic memory material is firstly registered into the episodic memory before being transferred into semantic knowledge. Semantic memory records knowledge in a categorical way (i.e. animals, plants etc.). Verbal fluency tasks, the vocabulary subtest of the WAIS and the Pyramids and Palm Trees tests are good examples of semantic memory tasks.

Attention, emotions, motivation and psychophysiological state interact with mnemonic processes. Therefore, before the assessment of explicit memory, accurate information about alertness, attention, motivation, and emotional and psychophysiological condition of the subject should be collected and evaluated.

A complete assessment of explicit memory needs of both episodic and semantic memory evaluations, but in some cases the assessment of episodic memory alone could be sufficient. These circumstances depend on the experimental protocol, the population studied and the goals of the study, and should be considered case by case.

Explicit memory can be considered an important outcome variable in assessing memory, but “per se” it is not sufficient for the substantiation of health claims in the context of increase or maintenance of memory.

Rey Auditory Verbal Learning Test

Rey Auditory Verbal Learning Test (RAVLT) consists of five presentations, with recall of a list of 15 words. The retention may be examined after 15-30 minutes, with distraction (i.e. visual spatial task) between the end of the presentations and the recall, or after hours or days. The examiner writes the words mentioned in the order in which the subject recalls them. If the subject fails to recall all 15 words, a recognition test will be administered in which 45 words are be verbally presented: the 15 words presented at the first trial plus 30 distractors(Vakil and Blachstein 1993).

The test includes seven trials (trials I-VII). Normative studies found that the immediate recall has a range between 6.3 and 7.8 words for people less than 70 years; the highest average (8.0 words) has been reported in a small group of university students.

To properly evaluate the performance of an examinee on this test, the deviation of the participants recall on each test session from the mean recall in that session for the normal population is calculated.

Although there are other validated tests to measure episodic memory (i.e. Logical Memory subtests of the WMS, Buschke Selective Reminding Test, the Paired Associates Learning subtest of the WMS, and other tests of verbal and visual recall and recognition) RAVLT is one of the most used, and it is suitable for quantitative assessment of both immediate and delayed recall capacities(Simard and van Reekum 1999). It is also appropriate to assess some interesting qualitative features of recall processes e.g. how the subject tackles the memory test, the serial position curve, the learning curve, and the assessment of possible intrusions and confabulations(Shah et al. 2014).

Furthermore, there are many validated alternative forms of RAVLT, so it can be used in both case/control and test/retest studies(Hawkins et al. 2004).

Like other neuropsychological tests, the validity and applicability of RAVLT must be evaluated in relation to the population under study.

Based on current evidence, RAVLT is appropriate for measuring episodic memory, although it cannot be considered a gold standard. For the evaluation of explicit memory, semantic memory (e.g. fluency tests) should also be assessed.

Fluency Tests

Fluency Tasks are a kind of psychological tests in which participants have to say as many words as possible from a category in a given time (usually 60 seconds). These categories can be semantic, such as animals, vegetables, clothes, etc., or phonemic, such as words that begin with a given letter. The semantic fluency test is sometimes described as the category fluency test. It seems that when a word or concept is activated in memory, and then spoken, it will activate other words or concepts which are related or semantically similar. This suggests that the order in which words are produced in the fluency task will provide an indirect measure of semantic distance between the items generated.

Data from this semantic version of the task have therefore been the subject of many studies aimed at uncovering the structure of semantic memory, in both normal development and mental illness. These tasks are widely used in neuropsychological practice for the assessment of semantic memory and/or language in functioning individuals(Ljungberg et al. 2013), as well as in various neurological disorders(Braga et al. 2009, Santos et al. 2014). In addition to measures of efficiency of search and selection of semantic and phonological/orthographic categories, these tasks also provide information regarding planning, organization, and cognitive flexibility. Accordingly, the functional imaging literature on verbal fluency demonstrated the involvement of a distributed brain network consisting of frontal, parietal, occipital-temporal area and anterior cingulate cortices in both tasks. Moreover, hippocampus has recently been shown to be involved in verbal fluency as well, particularly in tasks of semantic category (Glikmann-Johnston et al. 2015). This finding is in keeping with the idea that medial temporal lobe regions, and the hippocampus in particular, are involved in semantic memory.

Like other psychometric tests, the validity and applicability of Fluency Tests must be evaluated in relation to the population under study.

In conclusion, Fluency Tests, is appropriate for measuring semantic memory, although it cannot be considered a golden standard. For the evaluation of explicit memory, episodic memory (e.g. RAVLT) should also be assessed.

Implicit memory

Implicit memory is a type of memory in which previous experiences aid the performance of a task without consciousness of these previous experiences. Several studies confirm implicit and explicit memory as two separate entities. For example, patients with organic amnesia exhibit reduced explicit memory, but maintain largely intact their implicit memory. The most studied phenomenon in research on implicit learning and memory is repetition priming. Priming-related effects have been observed in numerous regions of the human brain by using fMRI and Positron Emission Tomography (PET). It was found that these brain regions depends on the type of stimulus and the manner in which it is processed (Reber 2013).

To evaluate the appropriateness of implicit memory as outcome variable of improvement, maintenance or reduced loss of memory, the literature deriving from database #4 was critically evaluated (Table 1).

The most studied phenomenon in research on implicit learning and memory is repetition priming. Repetition priming is a form of positive priming. This means that if a recently encountered stimulus is re-encountered, it is processed differently, usually more quickly. These behavioural effects reflect increased availability of previously seen items that is a memory trace dependent on the sensory cortex, and not on the MTL memory system. Functional neuroimaging studies of priming have consistently found that the neural signature of priming is a reduction in evoked activity for the re-encountered experience.

In conclusion, implicit memory can be considered an important outcome variable in assessing memory, but “per se” it is not sufficient for the substantiation of health claims in the context of increase or maintenance of memory.

Word Fragment Completion

Word Fragment Completion (WFC) is a test designed to measure the priming effect of implicit memory. Some words previously shown to participants are presented again in a fragmented form (i.e. missing letters) with the task of retrieving the missing letters from memory to complete the words. At the time of word presentation, participants have not consciously stored the items in memory; they have merely been exposed to them. In fact, to avoid participants consciously trying to retain the items presented, they were asked to perform tasks which require their conscious attention(Rajaram and Roediger 1993). Priming effect can then be observed when participants perform better on the WFC test for words that have been presented than for words that have not.

Other validated test (i.e. Word Stem Completion, Word Identification Task, Anagram Solution etc.) are often used to assess repetition priming(Soler et al. 2015). In all tests, priming effect depends on the typology of the stimulus with which the words are presented. Priming effect was greatest in verbal presentation (the subject listens and repeats the word), less in auditory presentation (the subject only listens to the word), and least in pictorial presentations. Moreover, it seems that each test has different efficacy in measuring priming effect depending on the type of stimulus presented. These facts highlight the importance of controlling the characteristics of the stimuli used when exploring the nature of priming.

WFC is a universally recognised test for measuring implicit memory throughout the assessment of priming effect. It has been showed to be independent of explicit memory and the correlation and consistency of this test with priming has been repeatedly evaluated.

Like other psychometric tests, the validity and applicability of WFCs must be evaluated for the population under study and the experimental conditions, on a case by case basis. In conclusion WFC is an appropriate choice for measuring implicit memory, although it cannot be considered a gold standard.

Conclusions

To date, several foods or food components have been proposed as subject of application for authorization of health claims related to cognitive functions, pursuant to Regulation (EC) 1924/2006. For most of them, EFSA has give out negative opinions due to the insufficient characterization of the food/food component, the choice of a not appropriate claimed effect, but mostly because of an insufficient substantiation of the claim. In this context, an effective substantiation of a claimed effect is provided considering many parameters affecting the quality of a randomized controlled trial (RCT), including an adequate choice of placebo/control, a proper sample size, and an adequate statistical analysis. However, the selection of adequate OVs and the related MMs used in the RCTs proved to be a crucial point when a certain health claim wishes to be associated to a food or a food component. The results provided by the present manuscript are relevant to drive the applicants towards a suitable choice of OVs and MMs in RCTs aimed at substantiating health claims on cognitive functions. Moreover, the results could help EFSA during the update of the guidance for the scientific requirements to bear health claims in the framework of cognitive functions. In addition to the use for health claim substantiation, the critical evaluation of OVs and MMs can be useful for the design of human intervention studies, and therefore can also impact general research.

Aknowledgments

This project has received financial support from the European Food Safety Authority (EFSA), Grant GP/EFSA/NUTRI/2014/01. The present article, however, is under the sole responsibility of the authors. The positions and opinions presented in this article are those of the authors alone and do not necessarily represent the views/any official position or scientific works of EFSA. To know about EFSA guidance documents and other scientific outputs of EFSA, please consult its website at: http://www.efsa.europa.eu.

Conflict of interest

The authors declare no conflict of interest.

References

Amodio D.M., Bartholow B.D., Ito T.A. Tracking the dynamics of the social brain: ERP approaches for social cognitive and affective neuroscience. *Soc. Cogn. Affect. Neurosci.* **9**: 385-393, 2014.

Appelle S., Oswald L.E. Simple reaction time as a function of alertness and prior mental activity. *Percept. Mot. Skills.* **38**: 1263-1268, 1974.

Arciuli J., Simpson I.C. Statistical learning is related to reading ability in children and adults. *Cogn. Sci.* **36**: 286-304, 2012.

Berch D.B., Krikorian R., Huha E.M. The Corsi block-tapping task: methodological and theoretical considerations. *Brain Cogn.* **38**: 317-338, 1998.

Bouman Z., Hendriks M.P., Kerkmeer M.C., Kessels R.P., Aldenkamp A.P. Confirmatory Factor Analysis of the Dutch Version of the Wechsler Memory Scale-Fourth Edition (WMS-IV-NL). *Arch. Clin. Neuropsychol.* **30**: 228-235, 2015.

Bouman Z., Hendriks M.P., Schmand B.A., Kessels R.P., Aldenkamp A.P. Indicators of suboptimal performance embedded in the Wechsler Memory Scale-Fourth

Edition (WMS-IV). *J. Clin. Exp. Neuropsychol.* **38**: 455-466, 2016.

Bowden S.C., Petrauskas V.M., Bardenhagen F.J., Meade C.E., Simpson L.C. Exploring the dimensionality of digit span. *Assessment.* **20**: 188-198, 2013.

Braga M.A., Tarzia O., Bergamaschi C.C., Santos F.A., Andrade E.D., Groppo F.C. Comparison of the effects of pilocarpine and cevimeline on salivary flow. *Int. J. Dent. Hyg.* **7**: 126-130, 2009.

Braun U., Schafer A., Walter H., Erk S., Romanczuk-Seiferth N., Haddad L., Schweiger J.I., Grimm O., Heinz A., Tost H., Meyer-Lindenberg A., Bassett D.S. Dynamic reconfiguration of frontal brain networks during executive cognition in humans. *Proc. Natl. Acad. Sci.* U S A. **112**: 11678-11683, 2015.

Brucki S.M., Nitrini R. Cancellation task in very low educated people. *Arch. Clin. Neuropsychol.* **23**: 139-147, 2008.

Brunetti R., Del Gatto C., Delogu F. eCorsi: implementation and testing of the Corsi block-tapping task for digital tablets. *Front. Psychol.* **5**: 939, 2014.

Bullinaria J.A. Modeling reading, spelling, and past tense learning with artificial neural networks. *Brain Lang.* **59**: 236-266, 1997.

Caban-Holt A., Abner E., Kryscio R.J., Crowley J.J., Schmitt F.A. Age-expanded normative data for the Ruff 2&7 Selective Attention Test: evaluating cognition in older males. *Clin. Neuropsychol.* **26**: 751-768, 2012.

Calvert G.A., Thesen T. Multisensory integration: methodological approaches and emerging principles in the human brain. *J. Physiol. Paris.* **98**: 191-205, 2004.

Cammisuli D.M., Innocenti A., Franzoni F., Pruneti C. Aerobic exercise effects upon cognition in Mild Cognitive Impairment: A systematic review of randomized controlled trials. *Arch. Ital. Biol.* **155**: 54-62, 2017.

Chechlacz M., Rotshtein P., Humphreys G.W. Neuronal substrates of Corsi Block span: Lesion symptom mapping analyses in relation to attentional competition and spatial bias. *Neuropsychologia.* **64**: 240-251, 2014.

Chelune G.J., Baer R.A. Developmental norms for the Wisconsin Card Sorting test. *J. Clin. Exp. Neuropsychol.* **8**: 219-228, 1986.

Chiu E.C., Koh C.L., Tsai C.Y., Lu W.S., Sheu C.F., Hsueh I.P., Hsieh C.L. Practice effects and test-re-test reliability of the Five Digit Test in patients with stroke over four serial assessments. *Brain Inj.* **28**: 1726-1733, 2014.

Coull J.T., Frith C.D., Frackowiak R.S., Grasby P.M. A fronto-parietal network for rapid visual information processing: a PET study of sustained attention and working memory. *Neuropsychologia.* **34**: 1085-1095, 1996.

Davis J.C., Bryan S., Li L.C., Best J.R., Hsu C.L., Gomez C., Vertes K.A., Liu-Ambrose T. Mobility and cognition are associated with wellbeing and health related quality of life among older adults: a cross-sectional analysis of the Vancouver Falls Prevention Cohort. *BMC Geriatr.* **15**: 75, 2015.

Drummond S.P., Bischoff-Grethe A., Dinges D.F., Ayalon L., Mednick S.C., Meloy M.J. The neural basis of the psychomotor vigilance task. *Sleep.* **28**: 1059-1068, 2005.

Eimer M. The neural basis of attentional control in visual search. *Trends Cogn. Sci.* **18**: 526-535, 2014.

Elwood R.W. The Wechsler Memory Scale-Revised: psychometric characteristics and clinical application. *Neuropsychol. Rev.* **2**: 179-201, 1991.

Eriksson J., Vogel E.K., Lansner A., Bergstrom F., Nyberg L. Neurocognitive Architecture of Working Memory. *Neuron.* **88**: 33-46, 2015.

Figueiro M.G., Sahin L., Wood B., Plitnick B. Light at Night and Measures of Alertness and Performance: Implications for Shift Workers. *Biol. Res. Nurs.* **18**: 90-100, 2016.

Fray P.J., Robbins T.W. CANTAB battery: proposed utility in neurotoxicology. *Neurotoxicol. Teratol.* **18**: 499-504, 1996.

Funke J. Complex problem solving: a case for complex cognition? *Cogn. Process.* **11**: 133-142, 2010.

Giesbrecht B., Sy J., Bundesen C., Kyllingsbaek S. A new perspective on the perceptual selectivity of attention under load. *Ann. NY Acad. Sci.* **1316**: 71-86, 2014.

Glikmann-Johnston Y., Oren N., Hendler T., Shapira-Lichter I. Distinct functional connectivity of the hippocampus during semantic and phonemic fluency. *Neuropsychologia.* **69**: 39-49, 2015.

Hartman D.E. Wechsler Adult Intelligence Scale IV (WAIS IV): return of the gold standard. *Appl. Neuropsychol.* **16**: 85-87, 2009.

Hawkins K.A., Dean D., Pearlson G.D. Alternative forms of the Rey Auditory Verbal Learning Test: a review. *Behav. Neurol.* **15**: 99-107, 2004.

Higgs S. Manipulations of attention during eating and their effects on later snack intake. *Appetite.* **92**: 287-294, 2015.

Kahn P.V., Walker T.M., Williams T.S., Cornblatt B.A., Mohs R.C., Keefe R.S. Standardizing the use of the Continuous Performance Test in schizophrenia research: a validation study. *Schizophr. Res.* **142**: 153-158, 2012.

Khitrov M.Y., Laxminarayan S., Thorsley D., Ramakrishnan S., Rajaraman S., Wesensten N.J., Reifman J. PC-PVT: a platform for psychomotor vigilance task testing, analysis, and prediction. *Behav. Res. Methods.* **46**: 140-147, 2014.

Kim H. Neural activity that predicts subsequent memory and forgetting: a meta-analysis of 74 fMRI studies. *Neuroimage.* **54**: 2446-2461, 2011.

Leverett J.P., Lassiter K.S., Buchanan G.M. Correlations for the Stroop Color and Word Test with measures of reading and language achievement. *Percept. Mot. Skills.* **94**: 459-466, 2002.

Ljungberg J.K., Hansson P., Andres P., Josefsson M., Nilsson L.G. A longitudinal study of memory advantages in bilinguals. *PLoS One.* **8**: e73029, 2013.

Lowery N., Ragland J.D., Gur R.C., Gur R.E., Moberg P.J. Normative data for the symbol cancellation test in young healthy adults. *Appl. Neuropsychol.* **11**: 218-221, 2004.

Mahoney J.R., Li P.C., Oh-Park M., Verghese J., Holtzer R. Multisensory integration across the senses in young and old adults. *Brain Res.* **1426**: 43-53, 2011.

Martini D., Angelino D., Cortelazzi C., Zavaroni I., Bedogni G., Musci M., Pruneti C., Passeri G., Ventura M., Galli D., Mirandola P., Vitale M., Dei Cas A., Bonadonna R.C., Di Nuzzo S., De Felici M.B., Del Rio D. Claimed Effects, Outcome Variables and Methods of Measurement for Health Claims Proposed Under European Community Regulation 1924/2006 in the Framework of Maintenance of Skin Function. *Nutrients.* **10**, 2017a. DOI: 10.3390/nu10010007.

Martini D., Biasini B., Rossi S., Zavaroni I., Bedogni G., Musci M., Pruneti C., Passeri G., Ventura M., Galli D., Mirandola P., Vitale M., Dei Cas A., Bonadonna R.C., Del Rio D. Claimed effects, outcome variables and methods of measurement for health claims on foods proposed under European Community Regulation 1924/2006 in the area of appetite ratings and weight management. *Int. J. Food Sci. Nutr.*1-21, 2017b. DOI: 10.1080/09637486.2017.1366433.

Martini D., Rossi S., Biasini B., Zavaroni I., Bedogni G., Musci M., Pruneti C., Passeri G., Ventura M., Di Nuzzo S., Galli D., Mirandola P., Vitale M., Dei Cas A., Bonadonna R.C., Del Rio D. Claimed effects, outcome variables and methods of measurement for health claims proposed under European Community Regulation 1924/2006 in the framework of protection against oxidative damage and cardiovascular health. *Nutr. Metab. Cardiovasc. Dis.* **27**: 473-503, 2017c.

Mcdowd J.M. An overview of attention: behavior and brain. *J. Neurol. Phys. Ther.* **31**: 98-103, 2007.

Mcfarland D.J. Modeling individual subtests of the WAIS IV with multiple latent factors. *PLoS One.* **8**: e74980, 2013.

Mulert C. Simultaneous EEG and fMRI: towards the characterization of structure and dynamics of brain networks. *Dialogues Clin. Neurosci.* **15**: 381-386, 2013.

Nyhus E., Barcelo F. The Wisconsin Card Sorting Test and the cognitive assessment of prefrontal executive functions: a critical update. *Brain Cogn.* **71**: 437-451, 2009.

Paiva G.C., Fialho M.B., Costa Dde S., Paula J.J. Ecological validity of the five digit test and the oral trails test. *Arq. Neuropsiquiatr.* **74**: 29-34, 2016.

Pearson K.E., Wadley V.G., Mcclure L.A., Shikany J.M., Unverzagt F.W., Judd S.E. Dietary patterns are associated with cognitive function in the REasons for Geographic And Racial Differences in Stroke (REGARDS) cohort. *J. Nutr. Sci.* **5**: e38, 2016.

Posner M.I. Measuring alertness. *Ann. N Y Acad. Sci.* **1129**: 193-199, 2008.

Purisch A.D. Misconceptions about the Luria-Nebraska Neuropsychological Battery. *NeuroRehabilitation.* **16**: 275-280, 2001.

Rajaram S., Roediger H.L., 3rd. Direct comparison of four implicit memory tests. *J. Exp. Psychol. Learn. Mem. Cogn.* **19**: 765-776, 1993.

Reber P.J. The neural basis of implicit learning and memory: a review of neuropsychological and neuroimaging research. *Neuropsychologia.* **51**: 2026-2042, 2013.

Rogers P.J., Hardman C.A. Food reward. What it is and how to measure it. *Appetite.* **90**: 1-15, 2015.

Ross S.A., Allen D.N., Goldstein G. Factor structure of the Halstead-Reitan neuropsychological battery: a review and integration. *Appl. Neuropsychol. Adult.* **20**: 120-135, 2013.

Sadeh A., Dan O., Bar-Haim Y. Online assessment of sustained attention following sleep restriction. *Sleep Med.* **12**: 257-261, 2011.

Santos I.M., Chiossi J.S., Soares A.D., Oliveira L.N., Chiari B.M. Phonological and semantic verbal fluency: a comparative study in hearing-impaired and normal-hearing people. *Codas.* **26**: 434-438, 2014.

Shah T., Verdile G., Sohrabi H., Campbell A., Putland E., Cheetham C., Dhaliwal S., Weinborn M., Maruff P., Darby D., Martins R.N. A combination of physical activity and computerized brain training improves verbal memory and increases cerebral glucose metabolism in the elderly. *Transl. Psychiatry*. **4**: e487, 2014.

Shalev L., Ben-Simon A., Mevorach C., Cohen Y., Tsal Y. Conjunctive Continuous Performance Task (CCPT)--a pure measure of sustained attention. *Neuropsychologia.* **49**: 2584-2591, 2011.

Simard M., Van Reekum R. Memory assessment in studies of cognition-enhancing drugs for Alzheimer's disease. *Drugs Aging.* **14**: 197-230, 1999.

Smith A., Sturgess W., Gallagher J. Effects of a low dose of caffeine given in different drinks on mood and performance. *Hum. Psychopharmacol.* **14**: 473-482, 1999.

Smith P.J., Blumenthal J.A. Dietary Factors and Cognitive Decline. *J. Prev. Alzheimers Dis.* **3**: 53-64, 2016.

Soler M.J., Dasi C., Ruiz J.C. Priming in word stem completion tasks: comparison with previous results in word fragment completion tasks. *Front. Psychol.* **6**: 1172, 2015.

Sternberg R.J. Intelligence. *Dialogues Clin. Neurosci.* **14**: 19-27, 2012.

Swanson H.L. Adults with reading disabilities: converting a meta-analysis to practice. *J. Learn. Disabil.* **45**: 17-30, 2012.

Vakil E., Blachstein H. Rey Auditory-Verbal Learning Test: structure analysis. *J. Clin. Psychol.* **49**: 883-890, 1993.

Van Horn J.D., Poldrack R.A. Functional MRI at the crossroads. *Int. J. Psychophysiol.* **73**: 3-9, 2009.

Weiss E.M., Siedentopf C., Golaszewski S., Mottaghy F.M., Hofer A., Kremser C., Felber S., Fleischhacker W.W. Brain activation patterns during a selective attention test--a functional MRI study in healthy volunteers and unmedicated patients during an acute episode of schizophrenia. *Psychiatry Res.* **154**: 31-40, 2007.

Yarkoni T., Speer N.K., Zacks J.M. Neural substrates of narrative comprehension and memory. *Neuroimage.* **41**: 1408-1425, 2008.

Zhao X., Yuan L., Feng L., Xi Y., Yu H., Ma W., Zhang D., Xiao R. Association of dietary intake and lifestyle pattern with mild cognitive impairment in the elderly. *J. Nutr. Health Aging.* **19**: 164-168, 2015.