Joachim Westenhoefer
France Bellisle
John E. Blundell
Jan de Vries
Damian Edwards
Wolfgang Kallus
Hubert Milon
Daphne Pannemans
Sandra Tuijtelaars
Hely Tuorila

Daphne Pannemans
Former ILSI Europe
Avenue E. Mounier, 83
Box 6
1200 Brussels, Belgium
E-Mail: publications@ilsieurope.be

Hely Tuorila
University of Helsinki
Department of Food Technology
P. O. Box 66
00014 Helsinki, Finland

Summary

Background The intake of food and drink can influence brain functions, which in turn may have effects on mental state and performance. Therefore, in principle claims to improve mood or specific aspects of cognitive performance by the consumption of functional foods are possible and indeed are currently found on the market. Aim The paper reviews existing methodologies, which may be used to substantiate and validate such claims of desirable effects of foods on mental state and performance. Results Mood, arousal, activation, vigilance, attention, sleep, motivation, effort, perception, memory and intelligence have been identified as relevant aspects of mental state and performance. The basic scientific concepts within this field as well as the methodologies to measure these concepts have been reviewed and described. Conclusions From this review it is concluded that, in principle, the phenomena in these fields are no different to those in other fields of life science. The scientific methods and protocols described in this report can positively demonstrate the effects of foods on mental state and performance in a scientifically valid way. A claim on mental state and performance like other claims must be based on scientific evidence. This report confirms that methodologies do exist to generate sound scientific evidence in this area. Therefore, claims on the enhancement of specific mental functions can and should be substantiated and validated using the methodologies described in this review.

Key words functional foods – brain functions – mental state – mental performance – mood
Introduction

It is an indisputable fact that food and drink put into the body can influence brain functioning. In turn, changes in brain function influence the output of that function, i.e. mental state and performance. Therefore, in principle the opportunity exists to develop foods and drinks that alter 'mood' or 'cognitive performance' in a way perceived to be desirable to consumers. It is obvious that mental state and performance can be adjusted over minutes, hours, days or weeks. In addition an effect may be generated immediately following ingestion but can be short lasting. Examples here would be the provision of glucose by a food, or the presence in food of some stimulant with a short half-life. Such effects could be instigated on each occasion that the food or drink is consumed. Alternatively, an effect may be induced more gradually and depend on the build up of key mediators in the body consequent upon continued consumption of the food or drink. Examples here are vitamins, minerals or essential fatty acids. Such nutrients could contribute to the structure, and function of biological systems including the brain. Such foods could help to maintain optimal functioning, enhance function or reduce factors leading to loss of function. Because of the wide range in the time course of the onset of action of effects on mental state and performance together with variability in the persistence of effects over time, claims in this area should make clear the latency to onset and the endurance of the action, in addition to the intensity of the response generated.

What follows in this paper is a scientific approach to claims in the area of mental state and performance. The paper seeks to clarify how to describe and measure the effects of nutrient intervention mediated by effects on the brain and related physiology. The subject has been classified in a way that best defines the component study disciplines comprising the body of knowledge in this area. It was decided not to include the study of satiety in this paper since this has already been covered in the FU-FOSE report [1], Dye & Blundell [2] and the COST-918 Report [3].

One purpose of this report is to encourage advertisers to express claims in specific rather than general terms. This is particularly important in the field of mental state and performance. This will create clarity of meaning for the consumer (and legislator) and, equally importantly, will permit a more decisive evaluation of claims. Accordingly it is proposed that claims should be worded in such a way as to be both measurable and quantifiable, and therefore to allow scientific verification. This proposal could restrict the creativity of some marketing departments but will permit a greater degree of scientific support for, and confidence in, the claim.

This paper will attempt to provide a framework for such debates with methodology proposals on how to validate such claims. Many manufacturers of food products, supplements or natural preparations have used on-pack claims to communicate the benefits of the product to the consumer. On a global basis various claims have been used on drinks, cereal bars and confectionery. Some example phrases seen on products and websites include:

- Stay active and alert
- Shortens reaction time
- Improve mental alertness
- Helps you sleep through the entire night
- Eases the tension caused by stress
- Has relaxing properties
- To calm and soothe
- Positive effect on mental sharpness
- Improvement of short-term memory
- Prevents memory loss

The above examples represent function specific claims because they refer more or less directly to one specific mental state, function or performance. The scientific concepts for these specific states and functions and the methodologies to measure them are reviewed in this paper. In addition to such function specific claims, other rather broad phrases have been identified on products or websites, as the following existing examples:

- Help you maintain peak performance
- With mind enhancing...
- Stimulate the body physically and mentally
- Is a brain food
- Support neurotransmitter activity
- Refreshing the cells of the brain
- Can help improve mental clarity
- When you need the maximum – physically and mentally
- Essential for mental energy, memory and focus

Such claims are designated as broad claims because they may not be linked to a specific scientific and measurable concept of mental state and performance or they do not specify a scientifically meaningful and valid way of how these concepts and processes are supposed to be influenced by the food or nutrient.

Please note that these qualifications are different from those spelled out in the ITG-D paper for “product specific claims” based primarily on human intervention studies with a particular food product, and “generic (food) claims” applicable to a range of products fulfilling certain criteria regarding compositions (The latter are referring to food products whereas in this paper the functional target is meant).
Outline of the paper

Following the introduction and some general considerations of methodology to assess the effect of food on mental state and performance, the present paper is organised into several sections, each section addressing a different area of mental state and performance. The areas covered within this paper are: 1) mood; 2) arousal, activation, vigilance, attention and sleep; 3) motivation and effort; 4) perception; 5) memory; and 6) intelligence.

Within each of these sections examples of potential claims related to the area of mental state and performance are listed, followed by a definition of the scientific constructs and concepts related to this field. After a short summary of potential mechanisms linking nutrition and diet to the specific mental function methods to assess it in a scientifically valid manner are reviewed. Finally considerations of potential problems in assessing the specific function and its relation to diet and nutrition are discussed.

Potential problems that are common to several of the mental processes are addressed in the concluding discussion of the paper.

General requirements

Generally, claims on functional effects of food and/or nutrients on mental state and performance have to be based on appropriate scientific evidence [4]. Such evidence has to be based on study designs, which allow to control as much as possible for potential biases and confounding. The accepted classification of evidence levels within evidence based medicine (EBM) [5] is not sufficient to study the enhancement of normal mental states and performances. Especially with regard to short-term effects on mental state and performance results from well-designed within-subject designs may represent an appropriate high level of scientific evidence. In such within-subject designed experiment one and the same subject is exposed to the different experimental conditions and repeated measurements of the dependent variables are taken.

Taking this into account, a modified classification scheme for the level of evidence is presented in Table 1. Claims of functional effects of food on mental state and performance should be based on evidence level I studies, on well-designed studies with evidence level II at the very least. Evidence level III is considered inappropriate to substantiate functional claims.

The measurement of mental state and performance is known to be susceptible to several experimental effects, as for example expectancies of the experimental subjects and/or expectancies of the experimenter. Therefore blinding of the experimental subjects and the experimenter with regard to the experimental condition (double-blind experimental design) may be crucial to yield valid results. To achieve blind conditions may be difficult, because functional ingredients may change the sensory properties of the experimental food samples and thus influence the experimental effects. For example in studying mood effects, functional ingredients may enhance bitterness [6] and thus, lead to an unfavourable response and even to an adverse effect on mood. Another example, in a recent study, respondents rating a beverage with an off-flavour did not consider it appropriate for maintenance of mental and emotional wellbeing (but were also, in general, negative to using a beverage for improvement of emotional well-being) [7]. In an experiment on mood effects, an unwanted sensory effect can be concealed by treating the control sample with another (bitter, if needed) substance to produce sensorically similar samples. Care must be taken to ensure that the interaction of the control substance with the food medium is similar to that of the experimental substance. Another example of masking sensory differences is the use of slight oral anaesthesia to prevent subjects from observing sensory differences between sucrose and saccharine [8]. This solution bears the problem of artificiality, in particular if the test product is to be used for prolonged periods of time.

In addition to the problems on sensory properties described above, a double-blind design is known to be a challenge in studies on food effects, in particular in cross-over designs which are preferred by some authors for good reasons [see 9]. Different control conditions may be worth considering, including e.g., serving noth-

Table 1 Levels of evidence for the evaluation of food effects on mental state and performance (modified from AHCPR 1992 [196])

<table>
<thead>
<tr>
<th>Level</th>
<th>Type of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ia</td>
<td>Evidence obtained from meta-analysis of level Ib studies</td>
</tr>
<tr>
<td>Ib</td>
<td>Evidence obtained from at least one randomised controlled study or from at least one well-designed within subjects – repeated measurement design</td>
</tr>
<tr>
<td>Iia</td>
<td>Evidence obtained from at least one well-designed controlled study without randomisation</td>
</tr>
<tr>
<td>Iib</td>
<td>Evidence obtained form at least one other type of well-designed quasi-experimental study</td>
</tr>
<tr>
<td>II</td>
<td>Evidence obtained from well-designed non-experimental descriptive studies, such as longitudinal observation studies, comparative studies, correlation studies and case-control studies</td>
</tr>
</tbody>
</table>
ing and serving water along a treatment with juice that contains a functional component [see 10]. In such cases, the subjects are likely to become aware of different treatments and enter the mode of impression management or social desirability bias.

Biomarkers in the assessment of mental state and performance

Enhanced cognitive function is a field where it is often easier to assess the function directly than measure any biomarker [2]. Although some physiological indices might appear good candidates as biomarkers of cognitive function (e.g. blood glucose level, changes in blood glucose level, measures of autonomic nervous system activity, heart rate, electroencephalography, etc.), they are more difficult and/or invasive to obtain than a specimen of the subject’s behaviour reflecting memory or attention at a given moment. Future research will certainly reveal certain strong relationships between biochemical or physiological parameters and cognitive performance. Such knowledge will contribute to the understanding of the specific effects exerted by certain foods or ingredients on certain aspects of attention or memory. It remains doubtful, however, that the monitoring of, say, blood glucose levels or cerebral activity over a given post-ingestive period will be more informative of beneficial effects than simply measuring cognitive performance in a variety of situations.

Biomarkers are essential for substantiating “Reduced risk of disease” claims. Valid biomarkers should represent intermediate endpoints of disease development. In terms of cognitive functions such as, e.g. memory (see below), signs of cognitive decline can be monitored in ageing subjects, before the eventual development of a full-scale disorder, such as dementia or Alzheimer’s disease. Studies carried out in ageing individuals suggest that certain parameters might be interesting biomarkers of the long-term effect of food choices and habitual diet on cognitive functions.

In the following sections, a number of validated tests will be proposed for the assessment of different mental states or functions. These tests address different aspects of mental activity. They cannot be ranked in terms of “accuracy” or “validity”, for example. The choice of one or many tests to assess a food’s or nutrient’s effects on mental functions should be guided by the type of mental change that is investigated, rather than by any inherent hierarchy in the value of the existing tests. All tests listed in this chapter have been used in earlier studies for assessing mental states. They are equivalent to other types of biological measurements and their results are tested using statistical methods. In many cases, since mental performance and behaviours are the results of complex interactions of mental operations, a battery of tests should optimally be used to support a claim.

Mood

Potential claims

- Improves your mood
- Improves your emotional well-being
- Helps to relax
- Helps to feel more energetic
- Makes you feel better
- Makes you happy/happier
- Helps avoiding sadness
- Helps avoiding depressive mood
- Reduces anxiety
- Improves depression

Definitions

Mood states in healthy population

Mood is an affective state based on or closely related to emotions [11, 12]. However, moods are considered less intense than emotions [13, 14]. Both moods and emotions are transient and variable, but moods may last minutes or hours, while emotions last only seconds or minutes [9]. A mood can be considered as a temporary predisposition to an emotion [14]. Typical mood states in healthy people are described as, e.g., happy, sad, energetic, calm, tense and anxious. Anxiety will be covered elsewhere in this review, in the section on motivation and effort.

Attempts to manipulate (non-pathological) mood states are often not targeted to a specific dimension of mood, but rather to general unspecified category of feeling states. In spite of the efforts to distinguish between emotions and moods, the same words and categories are often used interchangeably [see 11].

Improved mood is often treated as a mediating state for improved physical or mental performance [see 2]. There are also many studies measuring the effect of mood on the intake of foods [e.g. 15] or on responses to food or other stimuli [13]. Other studies deal with mood and physiological manipulations (such as tryptophan depletion) [16]. Thus, mood is often studied as an intermediate or intervening variable as well as an independent product or final goal in its own right.

Pathological mood states

A more specific, severe and long-term mood state is depression. Depression has been described as a state of reduced affect with symptoms of apathy, indifference, and...
loss of interest, or as an affective arousal with feelings of sadness, loss, guilt and despair, but confusion about the term exists [17]. A manic-depressive, or bipolar state is an affective disorder in which mood status fluctuates. Mania can be euphoric-grandiose or paranoid-destructive [18], both forms being characterised by increased activity, distractibility, poor judgement and other similar behavioural tendencies [17]. Affective disorders and their respective diagnostic criteria are detailed in the Diagnostic and Statistical Manual of Mental Disorders [19] or the International Classification of Diseases (ICD-10) [20].

Possible mechanisms

Mood is directly associated with specific functions of the central nervous system and thus, in order to affect mood state, a food or its component must influence the nervous system directly or indirectly.

A principal aspect of food and nutrition is that components enter the body via the digestive tract. Components influencing mood status either have to act in the GI tract or have to pass several barriers before becoming effective. A diversity of interactions between food and mood status can be identified. The stimuli generated by eating foods can be divided into neural (direct effect) and hormonal (delayed effect).

Upon entering the mouth, food may influence the mood by stimulating the sensory system or by prompting the responses associated with the cephalic phase of digestion. The effects on mood may depend on individual, learned preferences and on memory effects [12].

After entering the stomach, mechanoreceptors elaborate on the central nervous system, releasing hormones that prepare the intestine to digest the entering food [21]. It is imaginable that feelings of discomfort after a heavy meal, caused by mechanostimuli, affect mood status negatively. Also other mechanostimuli in the descending GI tract causing discomfort, such as gas formation in the colon, may affect mood. Equally it could be said that consumption of a large meal leads to a feeling of well-being, including relaxation, contentment and sleepiness. Perhaps more important than the action of food in the stomach, are effects (including CCK release) generated by the stimulation of chemoreceptors by the early entry of food into the upper small intestine. Animal research has shown that many gastrointestinal peptides act as neuromodulators which act on basic hypothalamic functions relevant to mood [22]. Not much research has been performed in this area.

If direct effects of food components on the central nervous system – specifically the limbic system in which mood regulation mainly resides – are to be expected, these components have to meet the following requirements. First, the foods have to be digested and components absorbed, thus crossing the mucosal barrier. After entering the mesenteric blood circulation these components have to pass the liver. There is a possibility that some components are transported by the lymph fluid and thus enter the main circulation without passing the liver circulation. The most selective barrier, however, is recognised in the blood-brain barrier. This barrier is highly permeable for water, oxygen, carbon dioxide and lipophilic substances. Other substances are selectively transported across this membrane [23]. However, this barrier is much less prominent in the hypothalamus, and in several small, specialised organs lining the third and fourth ventricles of the brain (the median eminence, area postrema, pineal gland subfornical organ and subcommissural organ). In addition, there is little evidence of a barrier between the circulation and the peripheral nervous system [24].

These physiological aspects should be taken into account when conducting studies on the effects of foods or nutrients on mood status. The methods that will be described below should therefore should mainly be regarded as black box models in which the real physiological properties of the interventions have not yet been elucidated.

Mechanisms of specific food components

Different food components can affect mood via different pathways. Thus, the effect of carbohydrates is believed to result from their capability of affecting serotonin levels by increasing the availability of tryptophan (precursor of serotonin) to the brain. More specifically, this effect is due to release of insulin, which increases uptake of large neutral amino acids into skeletal muscles, thereby removing the competition of tryptophan with those amino acids in the brain [1]. Individual variation in responses to carbohydrates exists [25].

Carbohydrates that have a high reward value, e.g., sucrose, may also prompt the release of endogenous opioids. For example, newborn babies suffering from pain can be calmed down by sucrose solutions [26]. Furthermore, the consumption of any highly preferred food may alter mood through the impact on serotonergic or opioid systems.

Caffeine acts as a competitive antagonist of adenosine receptors, causing a mild dilatation of the blood vessels, and increasing blood pressure, rennин and catecholamine release, urine output, CNS activity, metabolic rate, and intestinal peristalsis. Via its action on adenosine receptors caffeine affects activity of various neurotransmitters in the brain, including dopamine, serotonin, and noradrenaline [27]. The subjective effects are described as increased feelings of well-being, energy, motivation, self-confidence, alertness, stimulation, and concentration.
Methods (with examples)

Mood is a latent construct and therefore, it has to be measured using self-reports or the observation of behaviour. In the measurement of an individual's mood status, self-reported description of the emotions is used most commonly, although performance based assessment also has potential [9]. Scoring by a professional, rather than self-reporting, is used in clinical practice. A typical self-report consists of a questionnaire in which perceived mood characteristics are rated for their intensities or, less frequently, marked on a checklist of adjectives. There is a variety of validated adjective lists that have been used in the description of mood states [see 11].

Examples of instruments used for self-reports of mood

Some commonly used instruments to assess mood are presented in Table 2.

Profile of Mood States (POMS) has been developed by Lorr, McNair and coworkers [e.g. 28, 29]. POMS-BI [29] contains 72 adjectives to be rated on a 4-point scale according to what the feelings of a respondent are like (0 = much unlike this, 3 = much like this). Six subscales are based on these ratings. The subscales are composed–anxious, agreeable–hostile, elated–depressed, confident–unsure, energetic–tired, and clearheaded–confused. Examples of respective individual adjectives to be rated, selected from the positive side of the scale, are composed, friendly, cheerful, confident, attentive, and clearheaded. POMS scales have been applied in many studies related to food or nutrition, e.g., in [8, 25, 30–32].

Visual Analogue Mood Scales (VAMS) by Bond and Lader [33] consist of 16 bipolar scales, anchored verbally at each end of a 100 mm line. Similar to the mood rating scales by Folstein and Luria [34], these are often referred to as VAMS as well [e.g., 31]. In factor analysis, these scales reduce to three subscales of alerthness (9 items), contentedness (5 items) and calmness (2 items). Examples of respective bipolar scales are alert–drowsy, happy–sad, and calm–excited. Modifications of this device include a two-factor solution of these scales, complemented by two further visual analogue ratings, and resulting as separate scales for alertness and tranquillity [35]. The Bond and Lader scales and their modifications have been used e. g., by Spring et al. [31] and by Smith et al. [36].

Multiple Affect Adjective Check List (MAACL) measures the level of anxiety, depression and hostility [37] by allowing subjects to tick appropriate adjectives. A clinical validation study suggests that the scale reveals anxiety and depression well, but is less certain with respect to hostility [36, 37].

Beck Depression Inventory (BDI) is a 21-item rating inventory, most typically used as a self-report measure, tracking characteristic attitudes and symptoms of de-

Table 2 Commonly used tests to assess various aspects of mood

<table>
<thead>
<tr>
<th>Name or description of test</th>
<th>Target dimension of mood</th>
<th>Rating scale(s)</th>
<th>Developed and/or validated by</th>
</tr>
</thead>
<tbody>
<tr>
<td>Profile of mood states (POMS)</td>
<td>Energetic</td>
<td>72 adjectives rated from 0 = much unlike to 3 = much like this</td>
<td>[29]</td>
</tr>
<tr>
<td></td>
<td>Elated</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Agreeable</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Clearheaded</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Composed</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Confident</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visual analogue mood scales (VAMS)</td>
<td>Alert</td>
<td>16 bipolar scales rated on 100 mm VAS</td>
<td>[33]</td>
</tr>
<tr>
<td></td>
<td>Contended</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Calm</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiple affect adjective check list (MAACL)</td>
<td>Anxiety</td>
<td>Ticking appropriate adjectives</td>
<td>[37]</td>
</tr>
<tr>
<td></td>
<td>Depression</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hostility</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beck depression inventory (BDI)</td>
<td>Depression</td>
<td>21 items, for each 0 = normal, 3 = severe</td>
<td>[38]</td>
</tr>
<tr>
<td>Internal state scale (ISS)</td>
<td>Depression</td>
<td>17 statements rated on VAS scales</td>
<td>[39]</td>
</tr>
<tr>
<td></td>
<td>Well-being</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Perceived conflict</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Activation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hamilton depression rating scale*</td>
<td>Depression</td>
<td>21 items rated 0–2 or 0–4 for severity</td>
<td>[40]</td>
</tr>
<tr>
<td>Young mania rating scale (MRS)*</td>
<td>Mania</td>
<td>11 items rated on 0–4 or 0–8 for severity</td>
<td>[41]</td>
</tr>
</tbody>
</table>

* ratings by a clinician
pression [38]. The rated aspects of depression range from mood and sleep disturbance to feelings of pessimism, guilt and to loss of appetite and weight. Each aspect is rated on a 4-point scale (a statement at 0 reflecting no change from normal and a statement at 3 reflecting a severe disturbance). A score above 10 indicates mild to moderate depression, 19–29 indicates moderate to severe depression, 30–63 indicates severe depression.

**Internal State Scale (ISS)** tracks manic and depressive symptoms by self-ratings of 17-item visual analogue scales [39]. The subscales relate to depression, (lack of) well-being, perceived conflict and activation. Respective items, to be rated for severity and frequency, are, e.g., “I feel depressed”, “It seems like nothing will ever work out for me”, “I feel irritable”, and “I feel overactive”. The subscales on depression and well-being are well correlated with Hamilton’s [40] depression score, and the subscale on activation is correlated with the mania score by Young et al. [41].

**Instruments for scoring patients in clinical examinations**

**Hamilton Depression Rating Scale** is a 21-item scale that evaluates depressed mood, vegetative and cognitive symptoms of depression, and comorbid anxiety symptoms [40]. It provides ratings on current DSM-IV symptoms of depression (DSM-IV Manual of Psychiatrists [19]), with the exceptions of hypersomnia, increased appetite, and concentration/indecision. HAM-D score level of depression: 10–13 mild; 14–17 mild to moderate; >17 moderate to severe. This scale has been used in a nutrition-related study by, e.g., Wurtman et al. [42].

**Young Mania Rating Scale (MRS)** is an 11-item scale with a principle similar to the Hamilton [40] depression scale [41]. Each item, e.g., (elevated) mood, (lack of) sleep, (bizarre) appearance, is rated for severity on a 4-point scale, some of scales receiving a double weight, and the summed score is used for diagnosis.

**Other measurement practices**

In addition to multicomponent instruments, there are simple validated scales consisting of one single rating of mood. An example is the scale by Folstein and Luria [34] that utilises a visual analogue scale ranging from “my worst mood” to “my best mood”. The respondents mark their mood as it is “right now”.

Furthermore, rather than utilising the whole set of items, most researchers in the food area have used the instruments described above as a source to construct rating scales appropriate for testing specific hypothesis concerning effects of foods and food constituents [e.g., 10, 43–46]. Ratings of individual adjectives (e.g. not at all happy – extremely happy) and composite instruments, based on several pooled ratings with high correlations and a consequently high internal consistency have been used.

**Consideration of potential problems**

The initial mood and the status of a subject may have impact on the outcome. The initial non-pathological level of mood may be so high that it does not leave space for the potential improvement. The manipulation of the initial mood might be a solution in selected cases [47]. The initial status of mood is likely to show large inter- and intrasubject variations. Depending on the status of an individual, different responses may be triggered for example, stress-prone individuals may respond differently to a high-carbohydrate diet compared to less stress-prone subjects [25]. Furthermore, rather than being a true net benefit, the observed benefit may result from the recovery from deprivation of a substance (such as caffeine) from which a subject has been withdrawn [48].

To observe the consistency of effects in prolonged use, subjects should consume the product regularly. Consumers are known to respond negatively to monotony in sensory quality of food [49]. This could be overcome by offering a variety of options as vehicles of the functional ingredient. This, in turn, introduces another challenge, because this variety of acceptable flavour options needs to be developed. Furthermore, the particular flavours may be differently hedonically valued, and thus a further source of variation is introduced. Also, offering variety may be relatively easy with some products, such as beverages, but difficult with more complex solid foods.

**Arousal, activation, vigilance, attention, and sleep**

**Potential claims**

- Increases mental alertness
- Increases arousal
- Reduces daytime sleepiness
- Boosts mental speed
- Helps restful sleep
- Improves a good night’s sleep
- Improves sleep quality
- Enhances attention
- Enhances vigilance
- Enhances concentration
- Reduces drowsiness
- Reduces sleepiness
- Reduces somnolence
formation processing, from sensory feature analysis to the process of attention and detection of stimuli or information in the environment. In particular, the state of arousal is closely related to the sensitivities of the system to environmental events, and the magnitude and speed of response. Also included is the overall level of bodily activity and quiescence (known as waking and sleeping states).

Cognitive performance encompasses not only measures of speed (reaction time) but also of processing accuracy (measures of accurate and inaccurate detection). Interventions may elicit changes in function in some or all of these performance components. A failure to detect an effect of a particular macronutrient may be due to a true lack of effect, the effect of compensatory effort in experimental situations—"the Hawthorne effect" [50], or an effect on only a small component of the task (e.g., reaction time or decision time) such that an effect on overall performance is not detected. Recent advances in computer technology have meant that these tasks can be administered in a structured, carefully controlled fashion, with extremely accurate measurement of processing and reaction time, correct responses, and error rates.

Mental performance cannot always be neatly partitioned, and the state referred to as 'arousal' is integrated with processes of sensory detection, perceptual processing, and the detection of information in the environment. In particular, the state of arousal is closely related to the process of attention and detection of stimuli or events. Various tasks have been developed to assess various features of this detection process.

Attention and vigilance are major cognitive processes. They require a combination of all levels of information processing, from sensory feature analysis to the interpretation of the meaning of the stimuli. They have been extensively studied in humans as well as in animal species, and numerous test procedures have been developed to address one or many of the mental processes involved. Many recent reviews have described the effects of foods or food ingredients on these mental functions [e.g., 1, 2, 51, 52].

Attention is a state or activity of the brain predisposing the subjects to respond to some part or aspect of the environment rather than other parts. It represents the highest degree of conscience or awareness [53]. It is closely associated to arousal or vigilance, which are synonyms for wakefulness, alertness, and excitement.

Definitions

These dimensions of mental state and performance encompass those features concerned with the overall level of subjective and behavioral activity underlying the sensitivity of the system to environmental events, and the magnitude and speed of response. Also included is the overall level of bodily activity and quiescence (known as waking and sleeping states).

Cognitive performance encompasses not only measures of speed (reaction time) but also of processing accuracy (measures of accurate and inaccurate detection). Interventions may elicit changes in function in some or all of these performance components. A failure to detect an effect of a particular macronutrient may be due to a true lack of effect, the effect of compensatory effort in experimental situations—"the Hawthorne effect" [50], or an effect on only a small component of the task (e.g., reaction time or decision time) such that an effect on overall performance is not detected. Recent advances in computer technology have meant that these tasks can be administered in a structured, carefully controlled fashion, with extremely accurate measurement of processing and reaction time, correct responses, and error rates.

Mental performance cannot always be neatly partitioned, and the state referred to as 'arousal' is integrated with processes of sensory detection, perceptual processing, and the detection of information in the environment. In particular, the state of arousal is closely related to the process of attention and detection of stimuli or events. Various tasks have been developed to assess various features of this detection process.

Attention and vigilance are major cognitive processes. They require a combination of all levels of information processing, from sensory feature analysis to the interpretation of the meaning of the stimuli. They have been extensively studied in humans as well as in animal species, and numerous test procedures have been developed to address one or many of the mental processes involved. Many recent reviews have described the effects of foods or food ingredients on these mental functions [e.g., 1, 2, 51, 52].

Attention is a state or activity of the brain predisposing the subjects to respond to some part or aspect of the environment rather than other parts. It represents the highest degree of conscience or awareness [53]. It is closely associated to arousal or vigilance, which are synonyms for wakefulness, alertness, and excitement.

Possible mechanisms

The mechanisms involved in these processes ultimately depend upon brain neurochemical activity but clearly implicate intermediate biochemical events that intervene between the metabolism of the ingested food and the CNS neural activity.

Enhanced function claims

Attention is highly variable in humans and is sensitive to many factors such as the time of day (people are easier to distract in the afternoon than in the morning) that can easily obscure nutritional effects. Actually, most of the significant results published up to now [2] indicate deterioration rather than enhancement of attention following the intake of various foods or nutrients, suggesting that avoiding them might prevent a decline in cognitive performance.

Ingestion of fat at lunch and later can either disrupt attention [54] or improve it [55]. These effects are not entirely clear, however, since it appears that slower performance on attention tests can be compensated by higher accuracy [56]. High-CHO (high carbohydrate) lunches, which facilitate the selective uptake of tryptophan and serotonin synthesis by the brain [57], could play a decisive role in the "post-lunch dip", a phenomenon characterised by low attention capacity and somnolence in the early afternoon [31, 58]. The effects observed after manipulating the CHO:protein content of a meal are attributed to the induced depletion or enhancement of tryptophan [59]. High protein intake might increase the susceptibility to distracting stimuli [58].

Foods that could claim a beneficial effect on the "post-lunch dip", without exerting undesirable side effects, would certainly get much public attention. Since all three macronutrients have been associated with decreased attention, it is unlikely that such a claim could involve macronutrient composition. Other substances, such as caffeine might have a beneficial impact [60, 61]. Caffeine is a potent stimulant of the central nervous system. The effect of caffeine, however, depends on habitual intake and on various personal characteristics of the consumer [62, 63]. In addition, caffeine can exert adverse effects (increased nervousness, sleep difficulties, etc.).

The effect of foods or drinks on the sleeping state can be manifested as improved sleep quality itself or on the mental or performance benefits of having slept well. In fact, the effects of ‘quality’ sleep are more likely to be claimed as a subsequent benefit on other mood and cognitive parameters. As such, many of the methodologies discussed in these papers could be used to substantiate claims linked to having slept well with an appropriate study design.
In terms of effects of nutrients on sleep quality, there are several to list here. The effect of dietary tryptophan on serotonin levels has been discussed earlier in the paper and may be relevant to sleep quality too. The adverse effects of caffeine on sleep were also mentioned earlier, related to its antagonistic effect on adenosine, a calming agent for the cholinergic system [27]. Another dietary component that may affect sleep is glucose [64].

**Reduction of risk of disease claims**

Attention results from complex processes reflecting the activity of numerous brain structures. Efficient sustained attention is crucial for coping with daily life challenges. Any disruption of attention processes could have life-threatening effects. Under most circumstances, brain function is well protected by potent regulatory processes that maintain a stable output. The nervous system normally alternates phases of sleep, during which the organism rests, and phases of vigilance, during which the subject is conscious and attention can selectively focus on important stimuli.

Several processes could lead to more or less severe problems. Attention processes critically rely on the integrity of the nervous system. Inflammatory processes and vascular dysfunctions appear to play important roles in the pathogenesis of age-associated pathologies including Alzheimer's disease and dementia [52]. These processes could be critically affected by nutritional influences. For example, it has been recently proposed that both vitamins E and C are important for the central nervous system and that a decrease in their concentrations causes structural and functional damage to nervous cells. Abundant intake of these vitamins, from diets rich in fruits and vegetables, could lower the incidence of age-related neuro-degenerative diseases [52]. Dietary sources might be insufficient, however, to provide enough micronutrients to beneficially affect outcome in most individuals, and supplementation could be necessary to reach active amounts.

**Methods (with examples)**

**Biomarkers**

A number of procedures are now available for assessing reaction speed by direct measurement of brain physiological variables. One early recognised procedure involves the use of the EEG to measure evoked potentials. An early response detected by scalp electrodes is the so-called P300 waveform. The latency of the P300 has been proposed to vary with the length of perceptual processing but not with the duration of response selection. Therefore the P300 can be used to distinguish various stages in total response time.

**Measurement of subjective feelings of arousal, vigilance and alertness**

The subject's subjective impression of arousal, vigilance, alertness, sleepiness, and related sensations can be assessed using various psychophysical tools, among which the "Visual Analogue Scale" is a prominent example. Usually, it consists of a 100 mm long horizontal or vertical line, anchored at both ends with statements expressing extreme sensations, for example: “not sleepy at all” and “extremely sleepy”. The subject makes a pencil mark on the line to indicate the intensity of subjective sleepiness sensations at the time of the test [65]. An alternative is to use an adjective checklist in which subjects tick the sensations that apply at a particular moment in time.

**Measurement of the quality of sleep**

Methodologies that may be used to measure quality of sleep include those that measure sleep related changes in endocrine components such as histamine [66] or growth hormone [67], electrophysiological measures such as EEG (scalp), EMG (chin) and EOG (eyes) [68, 69] or even brain imaging techniques such as FMRI [70]. More common to use are questionnaires such as the Stanford Sleepiness Scale, a self-reported scale of treatment effect on daytime vigilance [71] or Leeds Sleep Evaluation Questionnaire, a subjective assessment of perceived changes in sleep following treatment [72]. Using tools such as these and those discussed elsewhere in this chapter it is possible to measure the effect of a dietary intervention on a subject's perceived quality of sleep.

**Vigilance monitors**

A recent study used ambulatory vigilance monitors that were custom-designed and manufactured. The monitors are lightweight devices somewhat larger than a wristwatch, which are worn on the nondominant wrist. They contain an 8-bit microprocessor, and various sensors that monitor, on a minute-by-minute basis, the subject's patterns of rest and activity. Among several other characteristics, the monitor contains a tone generator. It is programmed to assess vigilance by emitting an audible tone sequence at random intervals (about 20 times per hour). The subjects are required to respond to the tone by pushing one of the buttons on the monitor. The monitor records whether or not the subject responds to a particular tone signal, and the time required doing so [73].

**Critical Flicker Fusion Threshold**

The Critical Flicker Fusion Threshold (CFFT) is a means of measuring the ability to process discrete items of visual sensory data and is usually regarded as an index of central nervous system activity. A specific instrument is...
usually used to measure CFFT and is often incorporated into psychomotor testers. The instrument consists of four light emitting diodes placed in foveal fixation at a distance of one metre from the eyes. The lights flicker on and off systematically through a range of frequencies. The subject is required to discriminate flicker from fusion (as the frequency is increased – ascending direction) and fusion from flicker (as the frequency is decreased – descending direction), normally for three ascending and three descending trials. The CFFT is the point at which flicker is transformed into fusion and the psychophysical method of limits is used to calculate the mean CFFT. Numerous experiments on drugs have verified that the CFFT can be used as a measure of cortical activation. The mean CFFT correlates well with self-rated alertness using visual analogue rating scales [74]. CFFT is also positively correlated with EEG alpha activity and negatively correlated with reaction time [75], as well as with psychological test performance, fine motor movement, choice reaction time and duration of the spiral after effect.

Detection tasks

In these tasks subjects are usually required to respond to a particular stimulus or event by making a response such as pressing a key or a button. According to the type of reaction test (see later) the subject may be asked to make the same response to every stimulus or to vary the response according to the type of event. This approach is often based on the theoretical concept of a central executive or central resource model, which acts like a decision-making system. Indeed much of this work is based on a long-standing notion of signal detection theory. The task of the subject is to detect the presence of a signal against a background; the degree of difficulty of performing this task can be readily manipulated and is determined by a number of experimental features including the following.

1. **Number of stimuli:** Human operators are able to process between 5 and 9 ‘chunks’ of information at any one moment of time [76]. Therefore, task difficulty can be manipulated by systematically increasing the number of items, or ‘chunks of items’ presented to subjects, from 2 to 9 units.

2. **Stimulus presentation rate and time of recall:** The rate at which stimuli are presented influences performance accuracy, although the exact relationship is related to the particular task demands. For example, if subjects are asked to respond to each stimulus as soon as it is presented in a simple reaction time task, comfortable performance can be achieved at a rate of one stimulus/response per second – but task difficulty increases with faster rates of presentation.

   The relationship between rate of presentation and memory for items is less apparent. For instance, a high rate of presentation (or immediate recall of items) induces a short delay between encoding and retrieval, and can improve memory for sequentially presented lists. Alternatively, reducing the rate of presentation (or delayed recall) increases rehearsal time for each item, therefore improving memory.

3. **Serial versus simultaneous presentation of stimuli:** This is an important consideration for visually presented stimuli. For example, simultaneous presentation of a large number of stimuli may create an unnecessarily high visual load for subjects, taking their attention away from other visual tasks. Alternatively, serial presentation of visual information may be less demanding, especially if introduced at a slow rate, but if successful performance is contingent on remembering the order of item presentation, the task can become quite demanding.

4. **Types of tasks:** A number of psychometric tests have been used in published studies to assess the effects of foods or nutrients on various mental functions relevant to attention and vigilance (see Table 3). The following list briefly characterises these tests. It should be re-emphasised that these tests assess different aspects of mental performance and cannot be ranked in terms of accuracy and validity (see section Biomarkers above).

   **Dichotic shadowing:** Auditory stimuli are delivered to one ear while the subject is attempting to pay attention to stimuli delivered to the other ear [31].

   **Digit-symbol substitution task (DSST):** This test was originally developed as a component of the Wechsler Adult Intelligence Scale. Nine random three-row by three-column arrays of asterisks and dashes, labelled 1–9 from left to right are displayed on a computer monitor. A randomly generated number, between 1 and 9, displayed in the centre of the monitor, indicates which of the nine arrays should be reproduced on a three-row by three-column keypad. One response in each of the three rows, corresponding to the position of the asterisk, is required per trial and the third response per trial generates a new random number indicating the array to be reproduced on the next trial [32, 77–79].

   **Bimodal combi-test:** The subject has to concentrate simultaneously on two tasks. One task mobilises central short-term memory and consists of recognising defined sequences of coloured circles by pressing a button on a keyboard. These coloured circles are displayed in the middle of a monitor and appear one after another at the same place. In an additional peripheral attention task, the subject has to detect rotations of one of four pat-
terned circles that are displayed at the corners of the same monitor by pressing another button. Accuracy and efficiency scores are calculated for both tasks. A total score is also computed from subscores [55].

Bakan task (Rapid information processing task): This is a rapid visual information processing task with a high working memory load. Subjects are presented with a continuous sequence of single digits on the computer screen. The subjects press the space bar as fast as possible when they think they detect a sequence of either three odd or three even numbers. Presentation rate is 100 digits per minute with no inter-stimulus interval. Eight target sequences are presented each minute. The number of targets correctly detected and the number of false positives are recorded [44, 65, 80–83].

Repeated digits vigilance task: Subjects are shown three digit numbers on the screen at the rate of 100 per minute. Each number is normally different to the preceding one, but occasionally (eight times a minute) the same number is presented on successive trials. Subjects have to detect repeats and respond as quickly as possible. The number of targets correctly detected and the number of false positives are recorded, as well as reaction times [36, 56, 60].

Focused attention task: This test was developed by Broadbent et al. [84]. Subjects have to respond to a letter presented in the centre of a screen. Target letters are upper case As and Bs. On each trial, three warning crosses are presented on the screen to show where the stimuli will appear. Subjects are told to respond to the letter presented in the centre and ignore any distractors presented at the side. The crosses remain on the screen for 500 ms and are then replaced by the target letter. On some trials just the letter A or B is presented. On others, distracting stimuli (asterisks, letters) occur, one distractor being presented to the left of the target and another to the right. These distractors are presented close to the target or further away. Subjects have to respond by pressing different key either with the left (for letter A) or the right (for letter B) forefingers [36, 56].

Categoric search test: This test was also elaborated by Broadbent et al. [84]. Each trial starts with the appearance of two crosses in the positions occupied by the non-targets in the focused attention task. The subjects do not know which of the crosses will be followed by the target. The letter A or B is presented alone on half of the trials and is accompanied by a distractor digit on the other half. This task measures the importance of the presence of the distractor.
of distracting stimuli, the spatial separation of the two possible locations, and the tendency to orientate to the location in which the previous target was displayed [36, 56].

Sustained attention task: This test was developed by Neuchterlein et al. [85]. Single digits are displayed successively in the centre of a computer screen. The appearance of the digits is degraded by reversing the polarity of a random 30% of the pixels that define the digits and their background. Each digit is displayed for less than one second. Subjects are required to press the space bar each time a zero appears. The task goes on continuously for 10 minutes, during which a total of 600 randomly selected digits are presented, containing an average of 150 zeros [54].

Stroop test: This test developed by Stroop [86] deals with the analysis of conflict in sensory information. In one version of it, words describing colours are presented on a screen in different colours. At times, meaning and colour of the stimulus agree, at times they disagree. Subjects have to report colour of the stimulus, disregarding meaning, or vice versa [58, 82].

Hagen Central Incidental Test [87]: The test consists of six cards presented sequentially, each with a drawing of an animal and an object (e.g. a car). Once all the cards are presented, the subject is shown a single card with a picture of only an animal, and is asked to identify the serial position occupied by that animal in the first presentation of the series. At the beginning of the test, the subject is instructed to pay attention only to the animals, as the objects are incidental to the task. Two scores are generated: central (i.e. correct recall of animals) and incidental (correct recall of objects). This test was used in children by Pollit et al. [88, 89] in order to assess the cognitive effects of ingesting or omitting a high CHO breakfast.

Continuous attention test: Subjects are presented with a series of geometrical shapes and asked to respond whenever two identical shapes are shown consecutively [90]. The shapes consist of a four by four pattern of black and white squares. Stimuli are shown for 0.1 s and random interstimulus delays are of 2–4 s. The test involves 240 stimuli and the same pattern appears twice in a row on 40 occasions. The number of correct responses is recorded. A ceiling effect is avoided by presenting a secondary task: when a high pitched tone randomly occurs, subjects are required to press a left food pedal, when a low pitched tone is presented, subjects are asked to press a right food pedal [21].

Action response measurement

The measurement of the time taken to make a response appears to be a rather simple issue, but in practice there are numerous factors to consider. It has been proposed that when human subjects make a physical response our thinking about this can be organised according to whether the response is skill-based, rule-based or knowledge-based. However, in all cases the process involves the integration of perception, response selection and action. Knowledge-based responses normally form what can be called decision making tasks, and here the decision latencies are in the order of minutes, hours or days. However, for what we call reaction time measurement the decision latencies are in the order of milliseconds or seconds. Although accuracy is also monitored (see later), the latency of the response is the critical variable.

- Variables influencing RT. A large number of variables influence reaction time, one of the main ones being the degree of uncertainty. This variable differentiates the two main forms of reaction time. Taking the example of a sprinter in the blocks at the start of a race, there is a single stimulus (the sound of the gun), and a single response (leaping forward from the blocks). This defines the structure of the simple reaction time (SRT) task. Alternatively, a person driving a car may have to respond to an obstruction in the road at the left or the right, or to a person jumping in front of the vehicle. The appropriate response may be to swerve to the right or the left, or to brake. When there exists more than one possible stimulus (or event) and more than one response this defines the structure of the choice reaction time (CRT) task. In practice SRT tasks rarely occur outside the laboratory.

Other variables influencing RT concern the nature of the stimulus. For example stimulus modality; RT to an auditory stimulus is normally 30 to 50 ms faster than to a visual stimulus, and this is believed to be due to the speed of sensory processing. The intensity of the stimulus is also important particularly since this helps to improve the signal to noise ratio. Temporal uncertainty refers to the degree of unpredictability of the stimulus and therefore defines the extent to which the subject can be ready to respond. This factor can be modulated by introducing (and varying) a warning interval.

- Consideration of potential problems

Speed-accuracy tradeoff

A long known feature of response measurement is that CRT is longer than SRT; the main factor being the number of choices available. This issue refers to the reciprocity of latency and errors. Normally people make er-
errors as they try to respond faster. Subjects can be instructed to respond more accurately in which case they will be slower; alternatively, subjects can be told to respond fast in which case more errors will be made. Because the speed-accuracy ‘set’ can be shifted there is considerable interest in the ‘speed-accuracy operating characteristics’ (SAOC). This is particularly important when comparing RTs in different environments, laboratories and with different pieces of equipment. One particularly important feature of the SAOC is the compatibility of the stimulus and response, i.e. the extent to which the response is a natural (or contrived) reaction to the stimulus event.

Stages in reaction time

Models of CRT are often based on mechanisms of the human information processing system, and relationships exist between RT and the amount of information being transmitted (although this is often complex). The information theory model states that the total RT equals the sum of the durations of the number of component processing stages. This implies that data processing proceeds by discrete stages or through sequential mental operations. However, in the central nervous system it is recognised that a good deal of parallel processing occurs.

One approach envisages 4 stages: stimulus encoding, stimulus recognition (memory contact), response selection and response execution. The total RT will be the sum of the time taken for each of these separate processes. However, for both SRT and CRT tasks actual measurement can be simplified to separately compute processing time and movement time. This can be achieved, for example, by appropriate structuring of the experimental situation.

Response measurement in complex situations

One major consideration in measuring human performance is the complexity of the assessment task. In real environments it is very unusual for subjects to manage a single task without any distractors. More often a subject is required to perform on multiple tasks. These situations, often requiring attention to, and reaction to, a set of events going on against a background of other attention grabbing events are designed to measure the ‘functional state of the operator’. In the domain of functional foods it is clearly appropriate to think of foods influencing the operator functional state.

Consequently, the use of situations involving dual- or multi-tasking requires consideration. A number of assumptions need to be made concerning the use of performance tests to measure the effects of foods on operator functional state. The most important of these is that, because of regulatory control, performance may be protected by recruitment of effort, so that decrements may not be revealed in primary tasks. Therefore, measurement of performance must be accompanied by measurement of cost or effort of the psychophysiological state (see the section on motivation and effort below) in order to detect subtle changes such as a reduction in efficiency resulting from an increase in costs for measured performance.

Varying the workload (cognitive load) in tasks

In practice, a number of experimental tasks can be devised that require subjects to attend, and respond, to more than one element simultaneously. These tasks can be used to determine how the subject manages two sets of competing stimuli. Performance here is a function of the degree of arousal or the speed of information processing. This type of situation also includes complex tasks involving real life simulation including driving simulators in which subjects perform several functions simultaneously. The capacity to maintain performance under varying workloads depends on the degree of activation. Onset of sleepiness or a decline in alertness causes a decline in dual- or multi-tasking operations.

Rhythms in arousal: the post-lunch dip

One notable feature of the degree of arousal or alertness of individuals is that it displays a circadian rhythmicity [91]. The level of arousal fluctuates over the course of one day. These changes in arousal/sleepiness can lead to significant changes in work performance including an increase in errors, slowing of responses and delay in task completion. In turn this can lead to a decrease in productivity in work situations as well as constituting a hazard to health [92]. One noteworthy decrease in arousal and consequent decline in performance has come to be known as the ‘post-lunch dip’. The name derives from the fact that it is found in the early afternoon shortly after a mid-day meal. It is marked by a decrease in rated alertness and an increase in rated lethargy or sleepiness. Although this phenomenon occurs shortly after a meal, the consensus view is that it forms part of an endogenous rhythm and is only incidentally related to the prior meal [93]. However, it is likely that the size and composition of the meal do influence the intensity and duration of the ‘post-lunch dip’. Because of its proximity to a meal, the post-lunch dip represents a fascinating target for functional foods designed to combat lethargy. There may be a basis for a functional claim for a food or food component that can maintain performance at a uniform level across the middle of the day and prevent the decline often observed in many, but not all, people after the mid – day lunch. Indeed some individuals seem to be very vulnerable to this decline in arousal and performance. It may also be surmised that the concept of the
siesta – built into the cultural repertoire of certain societies – acknowledges the existence of the post-lunch dip and copes with it by incorporation into a legitimised period of resting. The post-lunch dip can be demonstrated by several procedures such as critical flicker fusion (CFF), visual detection and vigilance tasks that reflect the degree of central arousal, as well as by subjective ratings [94].

Motivation and effort

■ Potential claims

- Increases mental power
- Increases mental endurance
- Increases selective attention
- Increases goal orientation
- Increases frustration tolerance
- Counteracts fatigue
- Improves chronic fatigue syndrome
- Counteracts performance decrements
- Increases stress resistance
- Helps to sustain high performance
- Facilitates high performance
- Facilitates learning
- Facilitates effort
- Helps to cope with chronic fatigue syndrome
- Reduces depressive motivational states
- Counteracts burnout

■ Definitions

Motivation is a general concept, which explains start, direction and intensity of behaviour and cognitive processes like learning [95, 96]. Motivation is closely related to energetic consideration of behaviour and it is characterised by an affective component as it is closely related to pleasure and displeasure. Specific motivational states which refer to biological needs and states of internal homeostasis like hunger and thirst will not be covered in this section. Motivational components of emotional states are covered by claims on emotions. This section focuses on the energetic aspects of increased and sustained motivation. Increased motivation is accompanied by changes in arousal, which are due to voluntary efferent changes in central and peripheral activation, which are related to performance and goal related behaviour. Thus it is hardly possible to attribute performance changes either to activation or motivation unless in conditions in which these processes tend to dissociate like in deactivating conditions like fatigue.

“...The term motivation is used to mean the actual mobilisation of mental and behavioural effort in a particular encounter to achieve a goal or to prevent is thwarting” [96]. Incentive motivation and avoidance motivation can be distinguished due to the direction of behaviour. Incentive motivation will be the focus on this part, while avoidance is closely related to emotional states like anxiety [97].

Anxiety, depression and aggression are specific motivational states, which are treated effectively by psychopharmacological interventions. A large amount of knowledge has been accumulated on the neurochemical systems of these motivational states, which can be found in textbooks on the neurochemical basis of behaviour [e.g. 98]. These states are potential candidates of type B claims on motivational states. Chronic fatigue syndrome and burnout can be considered as further potential areas of type B claims.

These active changes of arousal are due to changes in effort. Effort has arousal components, is related to selective attention and is governed by higher central nervous processes, especially in tempo-frontal areas. Effort as well as motivation is inferred from changes in goal-oriented performance and/or activation in well-defined situations. Motivation and effort are often inferred from physical performance measures. Strength and speed of performance can be easily observed in motoric tasks like hand dynamometer or bicycle ergometer tasks. These parameters can be attributed to effect of food on effort and motivation as long as a direct metabolic effect can be ruled out. Thus claims on effort and motivation overlap with claims on endurance and physical performance as long as central nervous mechanisms are concerned. If CNS mechanisms can be assumed, claims can be substantiated by physical performance (cf. endurance) and mental performance paradigms as well.

As already stated by Dye and Blundell [2] counteracting performance decrements might be the more likely effects on cognitive functions than direct improvements.

■ Potential mechanisms

Motivation and effort are not restricted to specific areas of performance. The relation between changes in effort and motivation and changes in performance might differ markedly across different functions, as it is true for activation and stress. Simple reaction time tasks are improved by increased activation while complex tasks, which involve memory functions and complex problem solving are impaired by high levels of activation, stress and motivation. For complex tasks a speed accuracy trade-off might be the consequence.

Central nervous mechanisms, which mediate motivation, are linked to the mesolimbic dopaminergic system. This system plays an important role in motivation as it can be viewed as a central reward system [97, 99, 100]. This central dopaminergic system is involved in antici-
patory and consumatory aspects of motivation. Thus it can serve as a functional link between nutrition and motivational changes.

A second physiological model on motivated behaviour is related to the Henry and Stephen's two-process model of stress reactions [101]. This model claims positive effects on performance due to central and/or peripheral catecholaminergic activity, and performance-limiting effects of stress related corticoid activation of the hypothalamic-pituitary-adrenal system. The first system seems to be related more to an action oriented mental state, while the second systems is related to state orientation and anxiety related processes. These models predict a positive effect on mental and physical performance as long as nutritional ingredients help to energise behaviour via activation of central and/or peripheral catecholaminergic systems or help to provide adequate catecholaminergic resources. Depletion of catecholamines is related to performance decrements and depression and changes in mood state as well. Activation of the HPA axes will be accompanied by qualitative performance changes, which have been described in anxiety and stress research: focussing attention and restrictions in cue utilisation.

Effort and effort related changes in activation are closely related to goal oriented selective attention [102, 103]. Attention can be described within Baddeley's framework of working memory as a function of the central executive. Neurophysiologically attentional changes are related to frontal and perietal functions.

Both concepts allow establishing functional links between nutrition and performance changes.

Modern motivation theory stresses the role of higher cognitive processes like volition and decision-making. Especially decision making can be studied experimentally, thus experimental evidence about decision processes might substantiate motivational claims.

Methods (with examples)

Measurement of motivation

Direct measures of motivation in a given situation are speed, strength and persistence of performance. These parameters are easily assessed in physical performance (like hand dynamometer or bicycle ergometer performance) and possible for simple mental performance tests. Example from mental performance paradigms are

- Time until performance decrements appear in vigilance and monitoring tasks,
- Enhanced speed in self-paced sequential performance tasks (like mental arithmetic),
- Enhanced self-selected task difficulty in self selected item difficulty.

A weak but possible indicator of motivation and effort is a change in “energetic” arousal assessed by subtests on vigour or performance oriented activation in mood scales (POMS) or adjective check list, which assess mood multidimensionally. Classical methods to assess motivational states like Thematic Apperception Tests or Projective Tests need a complicated well controlled administration and a very high expertise to obtain reliable and valid data [104]. Projective tests will be the rare exception for substantiating claims, as the main validity of these tests concerns qualitative differences in motivational states and an appropriate training as well as high expertise is necessary for a proper interpretation of the results.

Motivation and effort can well be inferred from experimental studies.

- A first option to study motivational effect is given in experiments with performance hampering conditions, as motivation and effort allow to counteract fatigue, deactivation due to loss of sleep, post-lunch dip, time of day, alcohol, poor oxygen supply or inconvenient ambient conditions like heat or poor lightning.
- A second class of paradigms can be drawn from experiments, which affect motivation directly. Experiments on motivation like the Feather-Frustration Task [105], Paradigms to induce a state of unsuccessfulness and or state orientation. Stress induction by threatening self-esteem can be viewed as examples. Sustained performance in adverse conditions can be viewed as indicator for increased motivation and effort. Another motivational measure can be drawn from asking the subjects for rating of success probability. Other measures of motivational strength can be derived from adjusting task difficulty to the subject’s performance level by computer programs, which keep success-rate constant despite increased effort/motivation. Another way of affecting motivation is to look for changes of the effects of feedback (knowledge of results).
- A third group of paradigms can be based on experiments, which study the limitations of human information processing by double or multiple tasks. The double task paradigm is a basic method to check limits in information processing. Double task paradigms are used to measure the “spare” capacity, which can be used to increase effort. Tests on selective attention, which use distracting conditions like the Stroop Interference Test [86] are prominent tests for selective attention. It is recommended to use an approach which covers more than one level or paradigm to substantiate claims on effort.
- Finally experiments on higher cognitive functions, which involve simple and complex decision-making allow inferring on motivational processes. Decisions under uncertainty allow measuring changes in re-
response criteria and correctness of decisions via signal detection analysis. Repeated decision tasks allow to measure CNS activity (EP and ERD/ERS) [106] and thus give us objective information on changes in information processing.

**Biological indicators**

Using indicators of autonomic arousal for mental workload has been well established in work psychology [107].

Pupil size during task execution is a good candidate for a biological marker of motivation and effort, as it changes systematically with task difficulty. Another sensitive indicator is heart rate variability, which is derived from ECG recordings from the variability of the R-R intervals either by computing statistical variability measures (variance, standard deviation, mean of successive squares) or by the amplitude of low frequency components in fft-transformations (e.g. 0.1 Hz components). Decreased variability reflects rises in effort.

Finally EEG measures, which reflect selective attention, might also be used to substantiate claims of functional foods. Reliability and validity of the paradigm has to be insured in these instances as no standards are available yet.

**Examples in the area of nutrition**

A large number of studies with animals show clear-cut effects of nutrition (e.g. glucose) on performance and motivation related performance measures. This contrasts with a fairly small number of studies with humans, which address the notion of motivation/effort and nutrition directly. Studies either focus on mood or try to explain the performance change in an activation-oriented framework.

Some examples of studies, which might be considered a motivational framework, are listed in Table 4.

In addition, there are some examples of negative motivation states:

- **Fatigue** can be counteracted by caffeine and caffeine containing energy drinks (cf. Table 4). Long-term effects have been reported for fibre [108], which might be relevant to states like chronic fatigue. Effects of carbohydrate-rich and protein-poor diet on stress reactivity and feelings of depression have been reported by Markus et al. [25]. The authors claim serotonergic mechanisms as a possible pathway for the effects.

<table>
<thead>
<tr>
<th>Author(s) year</th>
<th>Paradigm</th>
<th>Measures</th>
<th>Nutritional variable</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benton, Slater &amp; Donohoe (2001) [155]</td>
<td>Different breakfast conditions</td>
<td>Time on test (recall)</td>
<td>Fasting</td>
<td>Fasting reduces time on test (effort)</td>
</tr>
<tr>
<td>Donohoe (2001) [198]</td>
<td>Sleep deprivation</td>
<td>Driving simulator performance, no explicit motivational measure</td>
<td>Energy drink (caffeine, taurine, glucoronolactone)</td>
<td>Drink compensates performance decline</td>
</tr>
<tr>
<td>Hovland (1936) [199]</td>
<td>Post-lunch dip</td>
<td>Physical energy expenditure in mental tasks</td>
<td>None</td>
<td>Amount of food increases effort during tasks</td>
</tr>
<tr>
<td>Lorist, Snel, Kok &amp; Mulder (1994) [200]</td>
<td>Prolonged work</td>
<td>Selective attention</td>
<td>Caffeine</td>
<td>Positive effect of caffeine in fatigued subjects</td>
</tr>
<tr>
<td>Frowein (1981) [Frowein, 1981]</td>
<td>Loss of sleep, prolonged work</td>
<td>Reaction time analysis</td>
<td>Amphetamines</td>
<td>Best effect in adverse conditions</td>
</tr>
<tr>
<td>Kennedy &amp; Scholey (2000) [140]</td>
<td>HR monitoring</td>
<td>Mental arithmetic, verbal fluency</td>
<td>Glucose drink</td>
<td>Better performance with “serial 7” accompanied by increased heart rate</td>
</tr>
<tr>
<td>Patat et al. (2000) [201]</td>
<td>Sleep deprivation</td>
<td>EEG, CRT, Stroop test and others</td>
<td>Caffeine</td>
<td>Increased performance on CRT, Stroop; Changes in EEG</td>
</tr>
<tr>
<td>van Baak &amp; Saris (2000) [202]</td>
<td>Bicycle ergometer and β-blockade</td>
<td>Bicycle ergometer – performance; FFA, Potassium</td>
<td>Caffeine</td>
<td>p = 0.056 – better endurance, less perceived exhaustion, no indication of peripheral effects</td>
</tr>
<tr>
<td>Seidl et al. (2000) [203]</td>
<td>Testing in the evening</td>
<td>ERP, d2, mood</td>
<td>Energy drink (caffeine, taurine, glucoronolactone)</td>
<td>Placebo ≥ decline prevented by energy drink</td>
</tr>
<tr>
<td>Alford et al. (2000) [204]</td>
<td>Anaerobic endurance, memory, performance measures (memory, concentration), HR</td>
<td></td>
<td>Energy drink</td>
<td>Significant effect of Energy drink on endurance, cognitive performance and heart rate</td>
</tr>
<tr>
<td>Nehlig (1999) [205]</td>
<td>Review</td>
<td>Dopamine</td>
<td></td>
<td>Reward system effects but not like amphetamines</td>
</tr>
</tbody>
</table>

---

Table 4  Studies on the effect of nutritional manipulations on motivation and effort
Consideration of potential problems

Performance increments must not necessarily be linked to increased effort. Increases in performance will also be observed following direct functional changes in specific functions, which contribute to performance (e.g. facilitation of signal detection) or due to changes in unspecific mechanisms like direct action on arousal. The close relation between effort and (psychophysiological) arousal is also shown in Kahneman’s definition. Recovery, which also counteracts performance decrements, should also be considered as possible alternatives to effort.

Measurement of effort can be based on performance measures or indicators of the underlying physiological or psychological mechanisms. Performance related claims on effort need a sophisticated experimental substantiation or specific paradigms due to the multitude of alternative mechanism, which allow sustaining performance levels despite fatigue and other performance decrementing factors.

Endurance in the physical area corresponds closely to the case of effort in the area of mental performance. For complex psychomotor functions, endurance and effort are functionally related. While endurance is primarily concerned with properties of muscle functioning, effort is primarily concerned with central nervous functioning.

The experimental set up or measurement paradigm has to make sure that the effect is not due to specific enhancement of performance levels to obtain an unobtrusive measure motivation and effort. Control conditions are necessary to separate non-effort related performance increments from effort-based increments.

Enhanced effort will show its effects especially in adverse conditions, while changes in performance levels will show its effects in optimal performance conditions. Factorial experiments are able to separate effects and test interactions. For example factor 1 could be caffeine vs. placebo, factor 2 will represent the adverse conditions (fatigued vs. not fatigued and “fatigue control”). If caffeine shows its effects independent of the performance modulation conditions, the claim should be “enhanced performance”. If caffeine acts most effective in the fatigued condition it will substantiate a claim of enhancing effort (or counteracting fatigue). To rule out a specific effect against fatigue, the effect should also appear for another performance adverse condition (best a more activating condition) like noise or distractions.

Definitions

Perception encompasses a set of related processes by which information from the environment is acquired and processed for integration into the cognitive system. It is convenient to distinguish three levels of perceptual processes: 1) At the sensory level physical stimuli from the environment are transformed by receptors into neural activity. The first steps of processing occur at this level by selecting and transforming sensory information for further processing. For example, processing of visual input in the retina emphasises limiting lines and luminance contrasts while continuous and homogeneous stimulation does not activate retinal output. 2) At the perceptual level information from the senses is organised and modified by superordinate processes in the brain into recognisable patterns. For example, three lines may be identified as the letter H or N depending on the context as for example the relation of positions of the three lines. These processes may result in estimates of sizes, motions or speeds of objects in the environment, which are dependent on internal computations, which in turn rely on formerly acquired knowledge, and its integration with the present information. 3) At the classification level the properties of the perceived objects are classified into appropriate categories. Round objects for example may be classified as either a football or an orange depending on context information, or objects may be identified as a face and further classified as “my mother’s face looking happy”. The perception processes encompass two sets of processes. The reception of sensory information by the senses and their feed forward to higher cognitive processes for the selection and analysis of relevant information represent bottom-up processes. The influences of knowledge, expectancies, memory and cultural background of a person on selection and interpretation of sensory input are designated as top-down processes. Usually perception encompasses an interaction of bottom-up and top-down processes.

Visual perception and acuity

In relation to dietary related effects, particularly visual perception has been the target of research. As a very basic index of visual perception acuity has been the focus of some research related for example to the effects of long chain polyunsaturate fatty acids on visual development in infants. Visual acuity is a measure of the smallest elements that can be resolved.

Perception

Potential claims

- Improve visual acuity
- Improves visual development

Potential claims

- See more clearly
- Improve night vision
Night vision

In the darkness the visual perception is vastly reduced. Perception of visual stimuli at low, scotopic light levels is mediated by the rod, or scotopic, visual system. Compared with the cone-mediated, or photopic, visual system, rod-mediated perception has poorer spatial and temporal resolution, poorer contrast sensitivity, reduced velocity perception of motions and totally absent colour perception [109]. A number of situations, e.g. steering a car at night, make it necessary to rely on rod vision and to cope with rapid changes between rod- and cone-mediated vision in order to estimate distances or the speed of moving objects. Night vision disturbances include glare disability (subjective reduction of visual perform-

■ Possible mechanisms

The role of long chain polyunsaturate fatty acids (LCP-UFA) on visual and cognitive development of infants has received considerable attention. A rational for this relationship lies in the observation that the retina and the brain are rich in omega-3 fatty acids, particularly in docosahexaenoic acid (DHA; 22:6n-3). DHA levels are particularly high in the membranes of photoreceptors and in the grey matter of the cerebral cortex [111]. Whereas human milk contains DHA and its precursor, alpha-linoleic acid (ALA, 18:3n-3), infant formulas used to contain only ALA, which is not easily converted to DHA by human infants [111]. Thus, studies found lower levels of DHA in plasma, erythrocytes, and the cerebral cortex of infants fed standard formulas as compared to breast-fed infants [112, 113]. Birch et al. [114] randomised preterm infants into three diet groups: 1) corn oil providing linoleic acid (LA), 2) soy oil providing LA acid and ALA and 3) soy oil supplemented with marine oil providing LA, ALA as well long chain omega-3 fatty acids. They found that only the marine oil supplemented group had visual acuity comparable to human milk fed infants whereas the other groups showed poorer acuity. In another study [115] full-term infants fed corn oil-based formula again showed poorer visual acuity than human milk fed infants.

Because the optical system of the eye is usually not the limiting factor of visual acuity in infants, the effects of fatty acids on acuity development could be attributed to changes in the retina or in central visual system [111]. In the retina the cone packing density continues to increase during the first 4 postnatal years. However, based on results in monkeys, these authors suggested that not the retina is the limiting factor, because the spatial solution measured in the visual cortex lags behind that predicted by the cone spacing in the retina. Thus, LCPUFAs and in particular omega-3 fatty acids may play a role in faster development and maturation of the infant’s brain. An other Cochrane Review of LCPUFA supplementation in infants born at term [117] concluded that there is little evidence from RCTs to support a benefit from LCPUFA supplementation for visual or general development.

■ Methods (with examples)

Inspection time task

Inspection time is considered to be a measure of perceptual speed and thus a measure of early information processing [118, 119]. In an inspection time task a visual stimulus is presented to the subject for a defined period of time and then visually masked [120]. A very simple stimulus may be used, usually two lines of markedly different length and the subject has to judge with line has been longer. Inspection time is taken as the shortest exposure duration at which a subject can make a correct judgement. Unlike reaction time, inspection time is not confounded by motor speed, motivation or the use of strategies [121].

Visual acuity

A number of standardised procedures to measure visual acuity exist. In adults acuity is usually assessed by reading a letter chart with rows of letters of different size, a procedure familiar to most people. One of the most widely used charts for visual acuity is the Snellen chart. However, a number of other charts exist which compensate some of the limitations of the Snellen chart [see 122 for further details].

In infants acuity may be assessed by using gratings consisting of black and white stripes or checkerboard patterns. Acuity can be measured using visual evoked potentials (VEP) or behavioural methods [116, 117].

Visual evoked potential is the electrical activity of the brain that is generated in response to reversing the black and white patterns of a strip of checkerboard grating. VEP is recorded from an electrode that is placed over the occipital pole. Usually response amplitude is linearly related to the size of gratings. A function generated can then be extrapolated to zero amplitude to yield an estimate of acuity threshold [111]. A newer method for VEP uses sinusoidal gratings, which are swept rapidly from low to high spatial frequency.
Behavioural methods to measure acuity in infants are based on the finding that infants strongly prefer patterned stimuli to non-patterned stimuli. In the forced preferential looking procedure an infant is shown two cards with different gratings (Teller acuity cards) and an observer views the infant through a peephole without knowing the density of the grating on the cards. He then makes a forced choice judgement about which card the infant prefers.

Vision-specific health-related quality of life

Impairment of visual perception and functioning may have significant impact on daily functioning, social activities and quality of life. In order to assess the subjective experiences of visual impairment a number of instruments have been developed. A recent review [123] identified 22 instruments assessing visual functioning and the impact of its impairment on health-related quality of life. The authors concluded that three of these instruments have been well validated and widely used: the Activities of Daily Vision Scale, the National Eye Institute Visual Function Questionnaire and the Visual Function Index.

Night vision

There exist several tests to assess night vision disturbances, i.e. decreased contrast sensitivity and glare disability. These tests are used to document and understand the effects of corneal refractive surgery. Fan-Paul et al. [110] give an overview and more detailed description of these tests together with a discussion of their advantages and disadvantages.

Dynamic vision at scotopic light levels has been investigated using flicker fusion (see section on arousal above). It has been shown that critical flicker fusion of the rod system is considerably lower than for the cone system [124]. To study motion perception at scotopic light levels Gegenfurtner et al. [109] used a sophisticated experimental setup displaying drifting sine-wave gratings on a computer-driven CRT monitor.

■ Consideration of potential problems

Precise definition of the target population for a claim is a crucial feature for claims on perceptual processes. As described above, nutrients may have different effects if administered to infants born preterm or born at full term. Moreover, effects may be present only during a limited period (some weeks or months) of time during cognitive and perceptual development. Because perception encompasses a broad spectrum of processes ranging from the basic sensory level up to the classification level, the specific type of perceptual process or property that is influenced by a food or drink has to be specified. Finally, the influence of food on perception may result from repeated intake of a nutrient over a period of time, as in the case of LCPUFA, or may be a short-term effect from a single intake of a nutrient.

Memory

■ Potential Claims

- Improves memory
- Improves short-term memory
- Improves long-term memory
- Reduces risk of memory decline

■ Definitions

Memory is a set of major cognitive processes that permit the short- or long-term storage of learned information, as a result of experience, and the recall of the acquired information (retrieval). Many recent reviews have described the effects of foods or food ingredients on these mental functions [for example: 1, 2, 51, 52].

Table 5 presents a list of various forms of memory, discriminated according to different dimensions. One easy discriminating dimension is the duration of the memory trace. Short-term memory, limited to a few minutes, is discriminated from long-term memory. Memory processes can also be discriminated according to the content of the remembered information: auditory or visual, spatial or verbal, etc.

Table 5  Forms of memory

<table>
<thead>
<tr>
<th>After Delacour [53]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) According to duration of retention: short-term or long-term memory</td>
</tr>
<tr>
<td>2) According to the type of content: auditory or visual, spatial or verbal, etc.</td>
</tr>
<tr>
<td>3) According to the type of cognitive process:</td>
</tr>
<tr>
<td>a) Non-declarative, or implicit memory, including Priming, Semantic, Procedural memory, Conditioning, Skills, Habits, versus</td>
</tr>
<tr>
<td>b) Strategic versus non strategic memory</td>
</tr>
<tr>
<td>Declarative, or explicit, memory, including Working memory (including three components): The central executive, The articulatory or phonological loop, The visuospatial scratchpad or sketchpad, Episodic, Semantic</td>
</tr>
</tbody>
</table>
The concept of working memory was elaborated by Baddeley [125] as a system for the temporary maintenance and manipulation of information, which allows the successful performance of complex mental functions such as comprehension, learning and reasoning. Baddeley’s model involves fractionating working memory into three sub-components: an attentional controller termed the central executive, aided by two active slave systems, the articulatory or phonological loop, which maintains speech-based information, and the visuospatial scratchpad or sketchpad, which holds and manipulates visuospatial information.

Long-term memory requires consolidation of the cognitive trace of experience. Several different forms of long-term memory can be defined according to the type of cognitive processes involved. The implicit versus explicit memory distinction is often used; it is also referred to as a procedural-declarative memory distinction. Implicit (or procedural) memory deals with learned or conditioned responses, such as habits and sensorimotor skills. It includes procedural memory of action chains. By contrast, declarative (or explicit) memory allows deliberate, voluntary recall of episodic or semantic information. Semantic memory refers to the storage of facts and general information, such as chemical formulas for example, while episodic memory has to do with recollection of personally experienced events. Strategic memory, according to Gabrieli [126] is an active process that mobilises many cognitive functions toward a complex goal, such as writing a book for example, whereas non-strategic memory is a passive process that elicits recall as a response to external stimulation. In certain cognitive disorders, various aspects of memory can be impaired while others are unaltered. Amnesic patients present a deficit of episodic memory but their implicit or procedural learning ability is intact. Patients with Alzheimer’s disease appear unable to update their semantic memory; their episodic memory is impaired for recent events while it can be excellent for long-past events.

Claims of enhanced memory should clearly define what type of function is affected.

**Mechanisms**

Memory performance is the result of complex processes reflecting the activity of numerous brain structures. Hence, several processes could lead to the more or less severe disruption of these cognitive functions. In healthy individuals, efficient cognitive performance is crucial for coping with daily life challenges. It is well protected by potent regulatory processes that maintain a stable output. Structures situated in the temporal lobe of the brain or in the diencephalon (medial thalamus, posterior thalamus) are involved in memory processes and a lesion of these structures induces a “global amnesic syndrome”, affecting declarative memory only, with a complete impossibility to form new memories. Interestingly, these syndromes are accounted for by attention deficits, since the structures most often involved are crucial in the regulation of the vigilance level [53].

There is a broad gap between normal, efficient cognitive function on the one hand and severe syndromes of dementia or amnesia on the other hand. Many mechanisms have been evoked to account for these variations in functional ability. A few examples of such mechanisms are briefly described below.

**Short-term effects**

- **Glucose supply to the brain.** In spite of the potent regulatory processes that maintain the adequate supply of fuel to the brain, in the form of glucose from the bloodstream, it has been argued that demanding cognitive tasks are facilitated when more glucose becomes available following the intake of pure glucose or a high CHO food. It is possible that demanding and/or prolonged mental tasks deplete blood glucose in certain areas of the brain, which then respond positively to the provision of an augmented supply of glucose [127]. This notion is supported by evidence from both brain scans [128] and monitoring changes in blood glucose levels during task performance [129]. Memory performance two hours after breakfast correlates with blood glucose level over a range of glucose values [130]. In agreement with this idea, it has been shown that young adults exhibited poor memory after missing breakfast, an effect that could be reversed by asking the subjects to drink a glucose beverage [131].

- **Glucoregulation capacity.** Glucose tolerance is assessed by a simple test involving the monitoring of the blood glucose changes following the ingesting of a glucose load. In healthy individuals, the blood glucose levels respond to ingestion of a CHO load (as a meal or a drink) by rising for about 30 minutes and then returning to baseline values within two hours. In persons with poor glucose tolerance, glucose reaches a relatively high value and then decreases slowly. In persons with type 2 diabetes mellitus, who have very poor glucoregulation capacities, cognitive impairments are sometimes reported, including selective attention deficits, verbal and visuo-spatial memory deficits [132–134]. These cognitive deficits are not permanent correlates of the condition, but can be reversed by short-term (1–2 months) treatment with oral hypoglycaemic agents [135, 136]. Normal variations in glucose tolerance could be associated with memory [83]. Better glucose tolerance, that is a faster and more efficient return to baseline blood glucose values after a meal, is associated with better performance on memory tests. In contrast, poor glucoregu-
lation is associated with poor cognitive performance following ingestion of a glucose load [137–140]. Impaired glucose tolerance has been identified as a predictor of cognitive impairment with age [141, 142].

- **Tryptophan.** Meals that are rich in CHO and poor in protein can theoretically affect the uptake of tryptophan by the brain and the synthesis of serotonin [57, see section on arousal above]. Effects observed after manipulating the CHO:protein content of meals have been attributed to the induced depletion or enhancement of tryptophan in the brain [59]. Tryptophan enhancement is associated with improved memory whereas tryptophan depletion has been shown to impair memory consolidation in healthy subjects [143].

**Long-term effects**

- **Antioxidant vitamin status.** The decline in memory capacity with age is a common observation. It is not known whether cognitive decline is an inevitable correlate of the ageing process or whether it could be prevented or retarded by appropriate measures, including nutritional interventions. A recent review [52] posits that inflammatory processes and vascular dysfunctions play important roles in the pathogenesis of age-associated pathologies, including dementia states such as Alzheimer’s disease. Human astrocytes, in particular, respond to CNS lesions by proliferation and cytodifferentiation. Astrocytes can produce a variety of cytokines, including interleukin-1 and tumour necrosis factor-α, as well as several adhesion molecules. These compounds appear to underlie pathologic processes and functional disturbances in acute and chronic neurologic disease. Taking generous amounts of both vitamins E and C, according to Martin et al. [52], may help to limit the build-up of cellular damage associated with disease and promote and maintain health and cognitive performance.

  Studies carried out in ageing individuals suggest that certain parameters might be interesting biomarkers of the long-term effect of food choices and habitual diet on cognitive functions. Plasma concentrations of vitamins (C and E for example) have been shown to correlate with cognitive impairment in dementia and Alzheimer’s disease [144–146]. Other studies have also suggested a close association between plasma levels of vitamins (E, C, beta-carotene) and better memory performance in people aged 50 and older [147, 148]. A review by Tangney [149] suggests that tissue levels of vitamin E, which reflect vitamin E intake, are associated with reduced cognitive decline with age. More work is needed before blood or tissue concentrations of certain elements of nutritional origin can be considered valid intermediate endpoints reflecting a level of specific disease risk.

- **Methods (with examples)**

  Memory functions are assessed by psychometric tests. Many tests exist for measuring different forms of memory (short-term, long-term, visual, spatial, verbal, non-verbal, etc.) (see Table 6). The following procedures have been used in investigations of food and nutrient effects. The references given in parentheses are examples of studies that used a particular procedure.

**Letters and verbal material**

- **Sternberg memory scanning task.** This test developed by Sternberg [150, 151] consists of four sub-tasks, corresponding to memory sets of three, four, five, and six different consonants. Each sub-task starts with a presentation of the particular set to be memorised on a computer screen, for as long as each subject needs. Then 60 individual letters are presented in succession for 1 second. Some belong to the memorised set, some do not. Subjects indicate as quickly as possible whether the probe letter does or does not belong to the set by pressing a red (No) or a green (Yes) button [152].

- **Sperling whole report task** is used as a sensory register task. Slides showing 12 random letters are presented on a screen during 50, 150, or 250 ms, in random order. Immediately after each flash, the subject has to report as many letters as possible within a response interval of 6 s [153].

- **Free word recall task.** Subjects read or listen to lists of words. Presentation rate of stimuli can vary (for example, one word per second). Subjects then have to recall as many words as possible. The number of words correctly recalled is recorded. Different types of lists can be used: semantic, acoustic, or neutral words [44, 81, 82, 130, 131, 153, 154]. Immediate and delayed (after an interval of a few minutes) recall can be assessed [83].

- **Free recall, old material.** Subjects have to report as many instances as possible of a certain predetermined word category (for example, fruits) in one minute [153].

- **Californian verbal learning test (CVLT).** This test was elaborated by Delis et al. [156]. It measures immediate, short delay, and long delay long-term memory (free recall; cued recall, and recognition) for a supra-span word list (for example 16 or 24 words) read at a fixed rate (one word every 2.5 s) [157, 158].

- **Delayed word recall task.** The same procedure as above is used, but a delay is introduced between the acquisition phase and the recall of words. Distraction can be exerted during the delay, using a variety of procedures, in order to prevent ceiling effects [155, 158].
<table>
<thead>
<tr>
<th>Name or description of test</th>
<th>Remembered material</th>
<th>Type of memory process</th>
<th>Developed/Validated by</th>
<th>Used in nutritional Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sternberg memory scanning test</td>
<td>Consonants</td>
<td>Recognition</td>
<td>[150, 151]</td>
<td>Adults (n = 19)</td>
</tr>
<tr>
<td>Sperling Whole Report Task</td>
<td>Letters</td>
<td>Immediate recall</td>
<td>[153]</td>
<td>Adults (n = 75), elderly (n = 23)</td>
</tr>
<tr>
<td>Free word recall task</td>
<td>Word lists</td>
<td>Immediate recall</td>
<td>[44, 81]</td>
<td>Adults (n = 18)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[130]</td>
<td>Adults (n = 16)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[154]</td>
<td>Adults (n = 33)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[82]</td>
<td>Women (n = 70)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[155]</td>
<td>Women (n = 150)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[153]</td>
<td>Adults (n = 75), elderly (n = 23)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[207]</td>
<td>Alzheimer’s disease patients (n = 23)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[208]</td>
<td>Down’s syndrome patients (n = 21)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[83]</td>
<td>Women (n = 46)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[56]</td>
<td>Adults (n = 48)</td>
</tr>
<tr>
<td>Free recall, old material</td>
<td>Subject’s own vocabulary</td>
<td>Cued recall</td>
<td>[153]</td>
<td>Adults (n = 75), elderly (n = 23)</td>
</tr>
<tr>
<td>Californian Verbal Learning Test</td>
<td>Supra-span word list</td>
<td>Recall (free, cued) &amp;</td>
<td>[156]</td>
<td>Adults (n = 65)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>recognition</td>
<td>[157]</td>
<td>Adults (n = 60)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Delayed recall</td>
<td>[155]</td>
<td>Women (n = 150)</td>
</tr>
<tr>
<td>Paried Associate Learning Task</td>
<td>Letters and flags</td>
<td>Associative recall</td>
<td>[153]</td>
<td>Adults (n = 75), elderly (n = 23)</td>
</tr>
<tr>
<td>Brown-Peterson Task</td>
<td>Consonant trigrams</td>
<td>Recall</td>
<td>[159]</td>
<td>Women (n = 80)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[160]</td>
<td>Women (n = 80)</td>
</tr>
<tr>
<td>East Boston Memory Test</td>
<td>Brief story</td>
<td>Immediate and</td>
<td>[162]</td>
<td>Elderly (n = 2889)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>delayed recall</td>
<td>[161]</td>
<td></td>
</tr>
<tr>
<td>Bakan Test</td>
<td>Digit sequences</td>
<td>Recognition</td>
<td>[82]</td>
<td>Women (n = 70)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[44]</td>
<td>Adults (n = 18)</td>
</tr>
<tr>
<td>Digit Span Test</td>
<td>Digits</td>
<td>Recall (forwards or</td>
<td>[164]</td>
<td>Children (n = 90)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>backwards)</td>
<td>[165]</td>
<td>Men (n = 16)</td>
</tr>
<tr>
<td>Number Recognition Task</td>
<td>Digits</td>
<td>Recognition</td>
<td>[166]</td>
<td>Adults (n = 12)</td>
</tr>
<tr>
<td>Serial Recall Task</td>
<td>Eight digit series</td>
<td>Recall</td>
<td>[167]</td>
<td>Adults (n = 144)</td>
</tr>
<tr>
<td>Running Memory Task</td>
<td>Digit series of unknown length</td>
<td>Recall</td>
<td>[167]</td>
<td>Adults (n = 144)</td>
</tr>
<tr>
<td>Spatial Memory</td>
<td>Drawings on grid</td>
<td>Recall of drawing</td>
<td>[131]</td>
<td>Adults (n = 33)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>positions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spatial and Temporal Memory Task</td>
<td>Sequences of eight lights</td>
<td>Recall sequence</td>
<td>[167]</td>
<td>Adults (n = 144)</td>
</tr>
<tr>
<td>Rey-Osterrieth Complex Figure Drawing Test</td>
<td>Complex geometric design</td>
<td>Immediate and</td>
<td>[168]</td>
<td>Adults (n = 60)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>delayed recall</td>
<td>[158]</td>
<td></td>
</tr>
<tr>
<td>Wechsler Memory Scale</td>
<td>Auditory, visual</td>
<td>Learning, Retention,</td>
<td>[170]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Retrieval</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logical Memory Scale (from Wechsler)</td>
<td>Short stories</td>
<td>Immediate and</td>
<td>[170]</td>
<td>Elderly (n = 260)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>delayed recall</td>
<td>[172]</td>
<td>Elderly (n = 137)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[173]</td>
<td>Adults (n = 184)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[131]</td>
<td>Alzheimer’s disease patients (n = 23)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[207]</td>
<td></td>
</tr>
<tr>
<td>Visual Reproduction Test (from Wechsler)</td>
<td>Geometric designs</td>
<td>Immediate and</td>
<td>[170]</td>
<td>Elderly (n = 137)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>delayed recall</td>
<td>[173]</td>
<td></td>
</tr>
<tr>
<td>Implicit Memory (Priming)</td>
<td>Pictures of familiar objects</td>
<td>Identification from</td>
<td>[148]</td>
<td>Elderly (n = 442)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>partial screen</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>presentation</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Paired associate learning task. Subjects have to learn ten combinations of letters and meaningless flags in as few trials as possible. Slides are then presented showing one flag. Subjects have to recall the particular letter associated with the flag [153].

Brown-Peterson task. This test, described by Peterson and Peterson [159], theoretically places great demand on the central executive of working memory, as defined by Baddeley [125]. Forty consonant trigrams are constructed (e.g. NQP), avoiding alphabetical sequences. At the time of testing, one trigram is spelled out and a three-digit number is given. The subject counts backwards in threes from this number at the rate of one calculation every 2 seconds. When a signal light flashes, after variable time intervals, the subjects attempt to recall the trigram [160].

The East Boston memory Test proposed by Albert [161] and Scherr et al. [162] involves immediate and delayed recall of a brief story [163].

Digits and numbers

The Bakan task (see section arousal and attention above). A series of digits are presented in rapid succession on a screen. Subjects must respond as quickly as possible by pressing a key when a sequence of three odd or three even digits is detected. The Bakan test is considered a test of attention but also has a high working memory load [44, 82].

Digit span test. Digits are presented to the subjects at varying rates. The subjects must remember (recall) the digits in forward or reverse order [164, 165].

Number recognition task. Delayed matching-to-sample trials are continuously presented on a computer [166]. Trials are signalled by a “Ready” massage. When subjects simultaneously hold down two keys, labelled “Y” and “N”, the “Ready” massage is replaced by a one to six-digit number, presented for 3 seconds. Two seconds later a test digit is presented. Within 2.5 seconds, subjects are required to lift the right index finger on the “Y” key if the test digit is contained within the original set of digits, and the left index finger holding down the “N” key if the test digit was not contained in the original set of digits. Accuracy and reaction time on this test reflect short-term memory [78].

Serial recall task. Eight single digits are consecutively presented on a screen at a rate of one per second. Subjects have to observe the eight numbers and then write them down in the order of appearance. When unsure of a number, they are encouraged to guess [167]. Serial recall of the order of digits provides a test of efficiency and rate of articulation, in Baddeley’s model.

Running memory task. This is similar to the serial recall task, except that the length of the digit sequence is unknown and subjects do not know when the sequence will end. This task was designed to examine the central executive of Baddeley’s working memory model [167].

Spatial stimuli

Spatial memory. Drawings representing different objects are placed on a grid and subjects are invited to concentrate for 20 s on the position of each picture on the grid. Then pictures are removed from the grid and subjects are asked to replace them on the grid, in random order, in their original positions. The time to finish the task and the number or errors is measured [131].

Spatial and temporal memory task. Subjects concentrate on five red buttons displayed on a response box. The red buttons light up in a randomised sequence consisting of eight lights. Subjects have to reproduce the se-

Table 6  Continued

<table>
<thead>
<tr>
<th>Name or description of test</th>
<th>Remembered material</th>
<th>Type of memory process</th>
<th>Developed/Validated by</th>
<th>Used in nutritional study by</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Working Memory</td>
<td>Pictures of familiar objects</td>
<td>Two simultaneous tasks: Identification from partial screen presentation Motor response to visual flash stimuli</td>
<td>[148]</td>
<td>Elderly (n = 442)</td>
<td></td>
</tr>
<tr>
<td>Explicit Memory</td>
<td>Picture containing objects, animals, digits, words, etc.</td>
<td>Recall &amp; recognition of picture elements after a 20 min delay</td>
<td>[148]</td>
<td>Elderly (n = 442)</td>
<td></td>
</tr>
<tr>
<td>Semantic Memory</td>
<td>Subject’s own vocabulary</td>
<td>Give definition of words</td>
<td>[170]</td>
<td>[148]</td>
<td>Elderly (n = 442)</td>
</tr>
</tbody>
</table>
quence by pressing the red buttons in the same order [167].

- Rey-Osterrieth Complex Figure Drawing Test. This task was developed by Osterrieth [168], after Rey [169], to evaluate long-term memory for nonverbal material. The subject is first asked to copy a complex geometric design and is then asked to draw it from memory, immediately and after a 30-minute delay [158].

- Implicit Memory Test. Pictures of familiar easy-to-name objects have to be identified when presented on a computer screen. In this procedure, pictures are generated on the computer screen by a recursive function that randomly selects a coordinate on the screen and from there on draws a line of 10 pixels of the picture. These lines continuously construct a picture that becomes clearer and clearer until it is completed. Priming effects can be assessed as percent gain in naming speed of known pictures compared with unknown pictures [148].

- Working Memory Capacity Test. This dual task involves the picture identification task just described, plus a simultaneous task requiring responding to light flashes presented either on the left or on the right of the screen by pressing on a key with the same-side index finger. Both picture naming and pressing in response to light stimuli have to be performed as quickly as possible. The dependent variables are the rate of correct responses to the flashes and the response time in the naming task [148].

- Explicit Memory Test. In an initial phase, subjects are presented with a picture on a computer screen. The picture contains two (left and right) almost identical scenes including objects, animals, digits, words, etc. Participants first scan for missing elements in the left-side image. Twenty minutes later, subjects are asked to recall freely as many elements of the scenes as they can remember. Finally, another picture containing elements of the previously scanned scene is presented and the subjects are asked to recognise these original elements [148].

Global memory capacity

In addition to changes in performance assessed using specific tests, it may be important to measure a person’s global memory capacity. This can be achieved using a number of multi-dimensional psychological tests, for example:

- Wechsler Memory Scale [170, 171]. This instrument assesses various aspects of memory functions (such as auditory immediate; visual immediate; immediate memory; auditory delayed; visual delayed; recognition delayed; general memory; working memory). In addition, tests of clinically meaningful aspects of memory functioning are evaluated (single-trial learning; learning slope; retention; retrieval). The test is administered in about 30–35 minutes and can be used in subjects’ 16–89 years old. The Logical Memory Scale and Visual Reproduction tests are drawn from the Wechsler test and assess verbal and nonverbal memory. The Logical Memory Scale requires subjects to listen to brief (one-paragraph) stories and attempt to recall them after presentation and again after a delay of 30 minutes. In the Visual Reproduction Test, subjects study each of four geometric designs and then attempt to draw the designs from memory, immediately and after a 30-minute delay [172, 173]. A Semantic Memory test has been derived from the Wechsler Adult Intelligence Scale Vocabulary test, and used by Perrig et al. [148]. This test requires subjects to give definitions of 32 words.

Consideration of potential problems

Using standardised tests of verbal memory presents a special difficulty. The above-mentioned tests dealing with verbal material were originally developed and validated in one language (most often English). Translating and adapting such tests to other languages require special validation procedures. Any translation of a psychological test from one language to another is a task that requires many professional skills to ensure that the original validity of the test is maintained in the translation. This is all the more delicate when the test is about language itself.

Defining the target population is an important aspect of enhanced memory claims. Immediate effects are observed most clearly in individuals with poor memory to start with, persons with borderline nutritional status, or varying degrees of cognitive dysfunction (Alzheimer’s disease patients, for example). In healthy individuals, it appears more difficult to improve memory, unless very demanding tests are used, probably because the brain is programmed for optimal functioning and protected from any momentary fluctuations in nutritional status. Enhanced memory claims should then be made for a clearly defined public. For most healthy people, in most situations, it may not be possible to improve memory by nutritional means.

Intelligence

- Potential claims
  - Increases intelligence
  - Helps to think faster
  - Helps to manage complexity
Definitions

Intelligence is an important dimension to describe differences between individuals and has a long history [174]. Although there is no generally agreed and undisputed definition of the concept of intelligence [175], many researchers and theorists may agree that intelligence covers the “ability to understand complex ideas, to adapt effectively to the environment, to learn from experience, to engage in various forms of reasoning, to overcome obstacles by taking thought” [176]. More specifically, many experts agree that intelligence covers two areas of abilities: verbal abilities and problem solving [177].

The notion of intelligence is closely related to the notion of cognition, and often these terms are considered as synonyms [178]. For example, Sparrow and Davis [178] refer to cognition as “the processes whereby individuals acquire knowledge from the environment” (p. 117) which closely resembles the previously cited concept of intelligence.

The first psychometric instrument to measure intelligence was introduced by Alfred Binet in 1905 in order to distinguish mentally retarded children from children of normal intelligence. The resulting test score was evaluated by comparing it to the average age of children achieving this score. This age was denominated as mental age. In adapting Binet's test for American children as the Stanford-Binet Intelligence test Terman (1916) used the concept of an intelligence quotient IQ, introduced by Stern (1914), which was calculated as

\[
\frac{\text{mental age}}{\text{chronological age}} \times 100 \quad [179].
\]

Nowadays, IQ is still used to describe the results from intelligence tests; however it is no longer based on the quotient of mental and chronological age, but on the statistical distribution of test scores. Usually, test scores are converted to an IQ scale with a mean of 100 and a standard deviation of 15.

The psychometric approach to intelligence has yielded a variety of intelligence tests (see below). Some of these tests consist of only one type of test items (e.g. matrices of patterns in Raven's Progressive Matrices), others include different types of items in different subtests (e.g. the Stanford-Binet or Wechsler Intelligence Tests). When using different subtests, individuals do not necessarily perform equally well on all different subtests. For example a subject may show relatively better performance in verbal comprehension than in complementing figures. Nevertheless, usually results on different subtests are positively correlated. When patterns of correlations between different intelligence subtests are analysed by factor analyses, a general factor g can be identified, explaining a considerable portion of common variance between the subtests [174]. Early on, theorists like Spearman [180] have emphasised the importance of \( g \) as a measure of general intelligence. In his view \( g \) represents some sort of hypothetical mental power, which underlies all intellectual operations. More recent approaches to this fundamental ability link \( g \) to differences in the speed of central processing of information [118], based of the finding that psychometric intelligence is correlated with the speed in very simple perceptual and cognitive tasks [118, 176]. The view that \( g \) represents a fundamental measure of intelligence is still maintained by many psychologists [e.g. 181], while others emphasise the importance of different abilities.

Cattell [182] introduced the concepts of fluid and crystallised intelligence. He hypothesised that \( g \) will be found not to be one factor but two. Crystallised intelligence \( g_c \) was described as “cognitive performances in which skilled judgement habits have become crystallised (whence it name) as the result of earlier learning application of some prior, more fundamental ability to these fields” [182]. Crystallised intelligence is seen for example in verbal tests or in achievements in geography or history. Fluid intelligence \( g_f \) is shown under circumstances that require the adaptation to new situations where crystallised skills do not offer an advantage. Typically performances at speed will be associated with fluid intelligence while performances without time limit, so called power tests, will be associated with crystallised intelligence. Cattell described fluid intelligence as directly physiologically determined whereas crystallised intelligence results from the experience of applying fluid abilities in a specific environment. Thus, crystallised intelligence is related to stored information and knowledge.

Guilford [183] developed a Structure of Intellect theory in which different factors of intelligence are classified according to their content, the required operations, and their products. The theory considers five different types of contents (visual, auditory, symbolic, semantic and behavioural), five types of operations (evaluation, convergent production, divergent production, memory, cognition) and 6 types of products (units, classes, relations, systems, transformations, implications). Each operation may be applied to each of the contents yielding each of the different products. Thus, a total of \( 5 \times 5 \times 6 = 120 \) specific factors of intelligence results. A specific factor would be, for example, the divergent production of semantic units. A test item for this factor could require subjects to list all items that he or she can think of that are round and hard.

More recently, Gardner [184] proposed a theory of “multiple intelligences”. In addition to familiar concepts

\[
\text{mental age} = \frac{\text{chronological age}}{\text{intelligence quotient IQ}} \times 100 \quad [179].
\]

Nowadays, IQ is still used to describe the results from intelligence tests; however it is no longer based on the quotient of mental and chronological age, but on the statistical distribution of test scores. Usually, test scores are converted to an IQ scale with a mean of 100 and a standard deviation of 15.

The psychometric approach to intelligence has yielded a variety of intelligence tests (see below). Some of these tests consist of only one type of test items (e.g. matrices of patterns in Raven's Progressive Matrices), others include different types of items in different subtests (e.g. the Stanford-Binet or Wechsler Intelligence Tests). When using different subtests, individuals do not necessarily perform equally well on all different subtests. For example a subject may show relatively better performance in verbal comprehension than in complementing figures. Nevertheless, usually results on different subtests are positively correlated. When patterns of correlations between different intelligence subtests are analysed by factor analyses, a general factor g can be identified, explaining a considerable portion of common variance between the subtests [174]. Early on, theorists like Spearman [180] have emphasised the importance of g as a measure of general intelligence. In his view g represents some sort of hypothetical mental power, which underlies all intellectual operations. More recent approaches to this fundamental ability link g to differences in the speed of central processing of information [118], based of the finding that psychometric intelligence is correlated with the speed in very simple perceptual and cognitive tasks [118, 176]. The view that g represents a fundamental measure of intelligence is still maintained by many psychologists [e.g. 181], while others emphasise the importance of different abilities.

Cattell [182] introduced the concepts of fluid and crystallised intelligence. He hypothesised that g will be found not to be one factor but two. Crystallised intelligence gc was described as “cognitive performances in which skilled judgement habits have become crystallised (whence it name) as the result of earlier learning application of some prior, more fundamental ability to these fields” [182]. Crystallised intelligence is seen for example in verbal tests or in achievements in geography or history. Fluid intelligence gf is shown under circumstances that require the adaptation to new situations where crystallised skills do not offer an advantage. Typically performances at speed will be associated with fluid intelligence while performances without time limit, so called power tests, will be associated with crystallised intelligence. Cattell described fluid intelligence as directly physiologically determined whereas crystallised intelligence results from the experience of applying fluid abilities in a specific environment. Thus, crystallised intelligence is related to stored information and knowledge.

Guilford [183] developed a Structure of Intellect theory in which different factors of intelligence are classified according to their content, the required operations, and their products. The theory considers five different types of contents (visual, auditory, symbolic, semantic and behavioural), five types of operations (evaluation, convergent production, divergent production, memory, cognition) and 6 types of products (units, classes, relations, systems, transformations, implications). Each operation may be applied to each of the contents yielding each of the different products. Thus, a total of \( 5 \times 5 \times 6 = 120 \) specific factors of intelligence results. A specific factor would be, for example, the divergent production of semantic units. A test item for this factor could require subjects to list all items that he or she can think of that are round and hard.

More recently, Gardner [184] proposed a theory of “multiple intelligences”. In addition to familiar concepts
like linguistic, logical-mathematical and spatial abilities he included musical, bodily-kinesthetic, inter-personal (understanding others) and intra-personal (understanding one self) abilities. One important point in Gardner’s view is that even in the areas of intelligence which are covered by traditional intelligence tests the often-used paper-and-pencil approach is not sufficient to measure intelligent behaviour which is relevant in everyday life, e.g. finding one’s way in an unknown town (spatial).

Another more recent theory of intelligence has been proposed by Sternberg [185]. This theory distinguishes the three fundamental aspects of analytic, creative and practical intelligence. Analytic problems, which are covered by traditional intelligence tests, are usually formulated by others, clearly defined, bring all necessary information with them, have only one right answer and little intrinsic interest. In contrast, practical problems often require problem recognition and formulation, are poorly defined, require information seeking, motivation and personal involvement and may have several acceptable solutions [176].

Carroll [186] reviewed the factor analytic literature on intelligence and produced as a summary a hierarchy of different abilities and factors with g at the apex. The abilities in this hierarchy below g are fluid intelligence, crystallised intelligence, general memory and learning, broad visual perception, broad auditory perception, broad retrieval ability, broad cognitive speediness, and processing speed. Each of these so-called stratum-II abilities in turn includes stratum-I factors, with may be either level factors (e.g. for fluid intelligence: induction, quantitative reasoning; for crystallised intelligence: verbal language comprehension, lexical knowledge) or speed factors (e.g. for broad cognitive speediness: rate of test taking, numerical facility; for processing speed: simple reaction time, choice reaction time). An Expert Task Force of the American Psychological Association concluded that the a factor hierarchy with g at the apex is currently the most widely accepted conception of the structure of abilities [176].

### Possible mechanisms

The effects of the diet and of specific nutrients on intelligence are typically considered to be long-term effects, which take periods of months to evolve.

Iron may play a critical role in cognitive performance. There is considerable evidence that iron deficiency anemia in children in developing countries is associated with poorer cognitive development [187]. However, the causal role of iron deficiency remains controversial. Results from randomised controlled trials support the idea that long-term iron supplementation of anemic children improves their cognitive abilities, although the evidence is not considered to be conclusive [187, 188]. Iron status could affect cognitive performance through several pathways including functional changes of the brain, alterations in the dopamine system, particularly the dopamine receptors or lags in the myelin formation, or delay of the acquisition of motor skills in anemic infants due to constraints of oxygen supply [189]. However, it remains unclear to what extent results on iron deficiency in developing countries can be generalised to industrialised countries.

The report of Benton and Roberts [66] that supplementation of British school children with a multivitamin-multimineral supplement resulted in improved non-verbal IQ as compared to placebo controls received considerable attention. Since then a number of studies have addressed this issue. In a recent review Benton [127] concluded that in 10 out of 13 studies, 12 being double-blind placebo-controlled randomised trials, at least a subset of children improved their non-verbal IQ as a result of vitamin and mineral supplementation. He argued that particularly children with a low supply of vitamins without being necessarily deficient might be responsive to supplementation. If supplementation resulted in an improvement of the brain chemistry it could be predicted that this should affect fluid intelligence, which is directly physiologically determined and hence non-verbal test scores, but not verbal scores, or crystallised intelligence. The results of the reviewed studies were in accordance with this prediction. Nevertheless, Benton admitted that was is no evidence to suggest a specific mechanism that underlies improved non-verbal intelligence. Thus, it is not clear, whether the supplementation resulted in an enhancement of normal function or in a repair effect of impaired mental functioning.

### Methods (with examples)

To measure the intelligence of individuals established intelligence tests have to be used. The development and validation of such tests as well as its application and evaluation has to meet several requirements which are documented in detail in the “Standards for Educational and Psychological Testing” jointly developed by the American Educational Research Association, the American Psychological Association and the National Council on Measurement in Education [190].

There exists a large number of intelligence tests, covering different subtests and therefore different mental abilities, designated for different age groups (children, adolescents, adults), to be administered individually or in a group format, and available in different languages. For example the publisher includes more than 40 different intelligence tests.

Table 1 summarises some of the more commonly
used intelligence tests many of which are available in several different languages. While the Wechsler intelligence tests are considered to be the most widely used Intelligence Tests [178], the Kaufman Tests have received considerable attention because they are theory based in contrast to the merely empirically developed Wechsler Tests. Especially for very young infants, up to 3:6 years the Bayley Scales are often used. They offer a Mental Development Index that reflects mental abilities and processes in early childhood and is considered as a measure of infant intelligence, although it does not measure the same abilities and functions that are assessed in intelligence tests for higher ages [189].

### Consideration of potential problems

The time and duration of dietary intervention, long-term vs. short-term, is important in the evaluation of effects on intelligence. Usually, tests of intelligence include subtests and/or items that are dependent on various cognitive processes and other aspects of cognitive performance, as for example arousal and activation or memory. Several of these aspects may be influenced by short-term manipulations of nutrients. This may create methodological problems, because it is possible that the measurement of intelligence is influenced by short-term manipulations of nutritional state, which is in contrast to the theoretical concept of intelligence as a relatively time-stable property.

The repeated administration of intelligence tests may represent another problem, because experience from the previous administration may result in learning or carry over effects. Therefore the use of control groups is essential. The use of parallel forms of intelligence test may be useful to minimise such carry over effects, if parallel forms exist.

### Discussion and general considerations

Although several indications exist of links between intake of foods or food ingredients and later improvements in mental state and performance, much remains to be done before any claim in this field can be considered to be truly substantiated. However, in principle, the phenomena in this field are no different to those in other fields of life sciences. This report has included a description of scientific methods and protocols that, if followed, can positively demonstrate the effects of foods on mental state and performance. It is argued that a claim on mental state and performance like other claims must

---

**Table 7** Commonly used intelligence tests (These tests are commercially available from the respective publisher)

<table>
<thead>
<tr>
<th>Test Name</th>
<th>Short name</th>
<th>Age group</th>
<th>Measures of intelligence</th>
<th>Individual or group test</th>
<th>Testing time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wechsler Intelligence Scale for Children – 3 rd ed.</td>
<td>WISC-III</td>
<td>6–16 years</td>
<td>Verbal IQ, Performance IQ, Full Scale IQ</td>
<td>Individual</td>
<td>60–90 min</td>
</tr>
<tr>
<td>Wechsler Adlt Intelligence Scale – 3 rd ed.</td>
<td>WAIS-III</td>
<td>16–89 years</td>
<td>Verbal IQ, Performance IQ, Full Scale IQ</td>
<td>Individual</td>
<td>60–90 min</td>
</tr>
<tr>
<td>Kaufman Assessment Battery for Children (K-ABC)</td>
<td>K-ABC</td>
<td>2:6–12:6 years</td>
<td>Sequential processing, Simultaneous processing, Mental processing, Achievement</td>
<td>Individual</td>
<td>35–90 min</td>
</tr>
<tr>
<td>Kaufman Adult and Adolescent Intelligence Test</td>
<td>KAIT</td>
<td>11–85 years</td>
<td>Crystallised IQ, Fluid IQ, Composite IQ</td>
<td>Individual (Core battery)</td>
<td>60 min (Core battery) or 90 min (Extended battery)</td>
</tr>
<tr>
<td>Stanford Binet – 4 th ed.</td>
<td>SB-4E</td>
<td>2–23 years</td>
<td>15 subtests scores, 4 area scores (Verbal Reasoning, Abstract/Visual Reasoning, Quantitative Reasoning, and Short-Term Memory), General score</td>
<td>Individual</td>
<td>35–85 min</td>
</tr>
<tr>
<td>Ravens Progressive Matrices</td>
<td></td>
<td></td>
<td>percentile rank of reasoning ability (Nonverbal)</td>
<td>Individual or Group</td>
<td>20–40 min (CPM, SPM) or 40–60 min (APM)</td>
</tr>
<tr>
<td>• Coloured Progressive Matrices</td>
<td>CPM</td>
<td>5–11 years</td>
<td></td>
<td>Individual or Group</td>
<td>20–40 min (CPM, SPM) or 40–60 min (APM)</td>
</tr>
<tr>
<td>• Standard Progressive Matrices</td>
<td>SPM</td>
<td>6–16 years, 18+ years</td>
<td></td>
<td>Individual or Group</td>
<td>20–40 min (CPM, SPM) or 40–60 min (APM)</td>
</tr>
<tr>
<td>• Advanced Progressive Matrices</td>
<td>APM</td>
<td>12–16 years, 18+ years</td>
<td></td>
<td>Individual or Group</td>
<td>20–40 min (CPM, SPM) or 40–60 min (APM)</td>
</tr>
<tr>
<td>Bayley Scales of Infant Development – 2nd ed.</td>
<td>Bayley-II</td>
<td>1–42 months</td>
<td>Mental Development Index</td>
<td>Individual</td>
<td>45 min</td>
</tr>
</tbody>
</table>
be based on scientific evidence [4] and this report confirms that methodologies do exist to generate scientific evidence in this area.

We reiterate the proposal made at the beginning of the present paper that claims have to be stated in function specific rather than broad terms. Claims should be worded in such a way as to be both measurable and quantifiable, and therefore allow scientific verification. This includes a wording that may be directly linked to the scientific concepts in the field of mood or cognitive performance as those that are reviewed in this report.

Generally, the mental functions covered in this report are compatible with the proposed interim criteria for the scientific substantiation of functional claims [4]. However, several issues require special consideration and remain to be resolved.

One particular issue is the amount of active substance required to induce a short-term or long-term effect. For example, the antioxidant vitamin doses required to induce a beneficial effect on changes in cognitive function with age are much higher than could be obtained from a well-balanced diet [191]. The amount of glucose needed to evoke a short-term improvement of memory for example is compatible with reasonable portion size. Ingesting pure sugar, however, does not appear a very practical means of influencing memory.

Claims should be substantiated for foods or ingredients as they are sold and consumed. The various effects attributed to glucose for example cannot be extrapolated to sucrose or other carbohydrates without adequate testing of the generalisation of effects.

The use of verbal scales as a methodological tool to assess aspects of mental state and performance poses particular problems, especially when these scales are translated and used cross-culturally. These problems are related both to the language and culture. Most of the scales have been developed in English, and the rest has to be translated into English to be introduced to the scientific community. Yet any self-administered questionnaires must be presented using the native language of the respondent and, furthermore, the rating scales used by professionals are usually translated to the first language of the professional before being used. The back-translation procedure – until the translated scale linguistically corresponds to the original – is currently a recommended practice to avoid a linguistic bias. However, this procedure does not omit the risk of culturally different interpretation [192]. The fundamentally similar or different interpretation is currently often sought using the latent variables in confirmatory factor analysis [e.g., 193]. Studying the ratings given in Beck’s Depression Inventory [38], Byrne and Campbell [194] found different item-response characteristics across cultures. Instruments that have not been cross-culturally validated need critical attention when translated and analysed, before firm conclusions can be drawn. Multimethod assessment may be the safest guard against measurement error or invalid interpretation [195].

Another methodological issue related to the validation of functional claims is the size of the effects of an ingredient on mental state and performance. As a general rule, statistical significance will not be sufficient to substantiate a functional benefit for the consumer. Even small effects without practical relevance may reach statistical significance if the sample size and thus the statistical power to detect small effects is large enough. To prevent such misleading conclusions from scientific research and subsequent misleading information of the consumer the effect size of functional effects has to be specified and to be discussed in relation to practical and/or clinical relevance. Functional claims should clearly be based on effect sizes that are of practical or clinical large significance.

Defining the target population is an important aspect of any claim related to improved mental state and performance. As discussed above, in the section on memory, some effects on mental performance may be present only in specific subpopulations, e.g. elderly people with poor memory or other cognitive dysfunctions. The target population definition may be as narrow as “preterm born infants during the first months of life” as this group has been found to profit in visual development from supplementation with long chain polyunsaturated fatty acids [116] whereas there is no evidence that infants born at term profit in the same way [117].

Nevertheless, the methodological considerations discussed in the present paper provide a framework for the scientifically valid investigation of possible functional effects of food on mental state and performance. Using these methodologies, claims of functional effects on mood and cognitive performance can and should be validated and substantiated.
References


76. Miller GA (1956) The Magical Number Seven, plus or minus two: some limits on our capacity for processing information. Psychol Rev 63:81–97


