Eur J Nutr (2008) 47[Suppl 3]:4–24
DOI 10.1007/s00394-008-3002-y

E. Isaacs
J. Oates

Nutrition and cognition: assessing cognitive abilities in children and young people

E. Isaacs
MRC Childhood Nutrition Research Centre, Institute of Child Health
University College London Institute of Child Health
30 Guilford Street
London WC1N 1EH, UK

J. Oates
Centre for Childhood, Development and Learning. Faculty of Education and Language Studies
The Open University, Walton Hall
Milton Keynes MK7 6AA, UK

ILSI Europe a.i.s.b.l. (✉)
Avenue E. Mounier 83, Box 6
1200 Brussels, Belgium
Fax: +32-2/762-0044
E-Mail: publications@ilsieurope.be

Abstract Although studying the effects of diet on cognitive function in children is of great interest to nutritionists, many have not received the formal training in the principles and practice of assessment necessary to properly evaluate research studies or to design and conduct research themselves. This paper is aimed at such an audience and assumes little prior knowledge of the field. A short description of neural development in childhood is followed by a discussion of important principles of assessment. A level of assessment approach is used to present a selection of widely used cognitive tests organized by cognitive domain and age group. Practical information about the tests is presented in tabular form and a list of useful websites is included.

Key words cognition – children – psychological assessment – nutrition

Introduction

Our aim in this paper is to provide information and guidance about cognitive testing to individuals, especially those working in the field of nutrition, who do not have a background in psychology or cognitive neuroscience. It may be that they wish to conduct studies to determine the effects of some nutritional intervention or they may simply want to be able to better interpret research reports or scientific articles. Our interest here is in the assessment of children; similar information for adults may be found in Schmitt et al. [34]. We have chosen to put certain constraints on the information we include in order to make the paper a manageable length. We consider here tests that are suitable for use when assessing effects of nutrition within a generally well-fed population consuming a Western-style diet. The reader should be aware, however, that these effects may be more subtle than those that would be found in children in certain developing countries. Studies of malnourished groups with marked deficiencies in energy, protein and some specific micronutrients may require a different approach and are, therefore, not considered here. Similarly, we discuss the assessment of cognition specifically, although mood and psychomotor function, requiring different forms of assessment, are commonly affected by nutrition.

Although the main intention is to give practical advice, we start with a short theoretical discussion about some aspects of cognition that are relevant to the nutritionist, followed by a few comments about the structure of cognition. We then describe factors relating to tests that need to be considered when designing a

EJN 3002

study, culminating in the choice of tests to be used. Of the many measures available, we have space to give examples only and the interested researcher is encouraged to look beyond the specific tests mentioned here, guided by principles that we will outline.

Cognition and nutrition

The cognitive structure that we measure in a child has come about through interaction between the brain and the environment over the course of development. In the initial stages of the development of the neural system in the fetus, the physical characteristics of the brain are determined by genetic factors but, immediately, the environment starts to influence this structure. Two fetuses with identical genetic make-up but exposed to different nutritional regimes during gestation, for example, may start to diverge in their neural development. After birth, the variety of environmental factors increases but nutrition continues to be important throughout life, particularly since it is now recognized that brain development occurs over a much longer portion of the life-span than originally thought [35]. Indeed, environmentally linked changes in brain architecture continue throughout the life-span: the brain is a very 'plastic' organ. Nutrition is a particularly important environmental variable to consider in relation to brain function because it can be manipulated relatively easily.

It is a general principle that the central nervous system is most vulnerable when it is developing rapidly. The time when nutrition has the greatest effect on brain development, therefore, is during the perinatal period known as the "growth spurt", considered to include the third prenatal trimester and the first few months of postnatal life in human infants [12]. Within this period of rapid growth, neural events occur according to a well-established timetable [8] so that the effects of nutrition will depend, to some extent, on when they take place. For example, during the first months of human gestation, the brain cells that are being produced are almost all neurons whilst, after 25 weeks, glial cells predominate [20]. Other processes such as myelination and synaptogenesis follow in order but it is important to note that the precise timing of these may differ for different regions of the brain. Determining the potential effects of nutrition must thus take into account the processes that are occurring at a particular time point in different areas of the brain. It has been suggested [15] that further spurts in brain growth in humans occur between 2–4, 6–8, 10–12 and 14–16 years. These may also prove to be times when nutritional interventions are effective; during the adolescent spurt, for example, myelination in the frontal lobe in particular is taking place [29] and nutrition might influence this process. The above all describe possible effects on the macrostructure and microstructure of brain anatomy but nutrition may also have a role to play in brain physiology by affecting both the level and operation of various neurotransmitters. To be clear, the influence of nutrition does not stop when the initial growth spurt ends. Neural growth and development continue throughout childhood and adolescence and need specific nutritional resources in sufficient quantities in order to occur and reach their maximum potential.

As well as affecting the way the brain develops, nutrition plays a crucial role in maintaining brain function. Greenwood et al. [18] point out that there are at least three important ways in which diet may affect neurochemistry. First, the ingestion of food affects the availability of the precursors required for the synthesis of neurotransmitters. Second, food serves as the source of the vitamins and minerals that are essential co-factors for the enzymes that synthesize neurotransmitters. Third, dietary fats alter the composition of the nerve cell membrane and myelin sheath, and that, in turn, influences neuronal function. Another example is glucose. Since it is the main metabolic fuel of the brain, the rate of glucose delivery from food to the blood stream and, thereby, the glycemic characteristics of the carbohydrate in the diet, could influence cognitive function as well [5]. It should be clear that not only severe malnutrition but also variations in the normal diet can influence neuronal function and hence cognition.

We have emphasized the importance of knowing the stage of neural development when trying to determine the effects of nutrition on the brain. Timing is important in another sense as well. Any changes to the basic neural architecture brought about by nutrition are likely to be long-term. Because neural development is a cascade of events, even a small change early in the sequence may have a large and diffuse impact on future development.

It is also important to recognize that these early emerging differences in cognitive capacity linked with diet will, in turn, affect individuals' behavior and hence their selective use of the environment. Humans do not respond passively to the environments in which they find themselves; cognitive processes affect choices of environments and cognition is clearly implicated in dietary choices. These complex interactions among genetic, constitutional and environmental factors are usefully considered as 'transactional' processes [31].

Because nutritional interventions later in life are more likely to be short-term and specific, we might expect that early interventions would have widespread effects on overall cognitive level, reflected in IQ scores, while later ones would have less influence on such measures. Instead, they might affect performance in specific cognitive domains, like attention and memory. (Of course, early intervention may affect specific as well as more general domains of cognition, so it is important to test for these). The performance–competence distinction is highly relevant here; cognitive performance, reflected in behavior, is open to dietary effects at all times, whereas competence, more akin to underlying ability, is subject to developmental variations in susceptibility to dietary variations. The effects of eating breakfast on cognitive performance, for example, would be expected to be short-term (although possibly repeated daily) and might affect a specific function such as attention. Whenever they occur, nutritional effects are likely to be subtle since they will not result in gross distortions of brain structure. Rather than destructive lesions, we would expect to see "a quantitative and qualitative alteration of growth" [11]. As a result, we would not predict such marked cognitive changes associated with nutritional variance as we would in the case of neurological damage. This is important from an assessment point of view since many neuropsychological tests have been devised to detect brain damage of a sort we are not discussing here. These tests may not be sensitive enough to detect nutritional effects.

How a child performs on a cognitive test will also be affected by factors that are considered to be non-cognitive in themselves such as mood, motivation and effort. Performance may also be affected by the context in which the testing takes place: are there peers present? Is the adult familiar or a stranger? Is the setting home or school? In a study, one can usually control the context factors by keeping conditions the same for all participants. Factors such as mood are harder to control but if they are thought to vary widely among the children in the group, it may be advisable to obtain some objective measure that can be used in the analyses of the data. It is well established that socio-economic factors have strong effects on the development of cognitive competence, through a variety of influences including parenting, availability of resources in the home and access to educational opportunities. Given the multiplicity of transacting factors influencing the development of cognitive abilities, teasing out independent effects of nutrition presents a daunting methodological challenge and requires careful experimental design.

The structure of cognitive functions

We use the term "cognitive functions" to refer to the activities of taking in information from the environment, processing this internally and then responding in the form of behavior. We can think of cognitive functions as forming a hierarchy, moving from the general to the specific.

The most general level is that of overall intellectual ability, usually described as intelligence. There are two reasons why there has been so much debate within psychology about the existence of a general intelligence factor, often referred to as "g". First, it is an empirical fact that performance differences between individuals (both children and adults) in different cognitive domains covary; people who are verbally fluent tend also to be good at mental arithmetic, for example. This suggests the presence of overall or global general intelligence differences between people. But these covariations are not perfect; each individual has a unique 'profile' of strengths and weaknesses in different cognitive domains, doing better in some areas than others. So it is not possible to simply rank people on a general intelligence scale. Some psychologists [23] think that performance is determined by a general factor and a series of more specific factors relevant to certain, but not all, tasks. It may be that "g" will turn out to describe some aspect of the nervous system that affects all cognitive performance, such as speed of nerve conduction, but this factor has not yet been described (see [9] for discussion of this whole topic).

The second reason for debate is that the idea of "g" has often been linked with a notion of genetically determined differences in intelligence between individuals, and, more politically contentiously, with genetically determined differences in intelligence levels between groups. Without going deeply into that debate, it is worth commenting that it is now becoming clear that virtually all human abilities and psychological characteristics are influenced to a degree by genetic factors, but that environments are also powerful influences in the transactional processes referred to above. Increasingly, gene $\times$ nutrient interactions are being discovered. Caspi et al. [7], for example, have recently reported that genetic variation in fatty acid metabolism moderates the effects of breastfeeding on cognitive development, measured by IQ scores.

It is the concept of "g" that has led to the development of the well-known measure "IQ", standing for "intelligence quotient" and based on a measure called mental age. Originally, intellectual level was defined as the ratio of a person's mental age relative to their chronological age; the formula to calculate it is MA/CA $\times$ 100. Thus, someone whose mental age is 12 with a chronological age of 10 will have an IQ of 120. Someone whose mental age is the same as their chronological age will have an IQ of 100. Although modern tests derive IQ in a slightly different way, they are still based on the idea of comparing an individual's performance to that of a large group with the same chronological age. Chronological age can be measured in a valid and reliable way as long as we know accurately a person's date of birth (except in the case of premature birth, where corrections are usually made for the first 2 years). However, mental age is a much more problematic measure, since it depends also on knowing what the mean performance levels are for individuals of varying ages across the population from which the person comes. It also makes assumptions about the shape of the distribution of such scores from the test being used to assess IQ. For these reasons and because of the many other sources of possible variation in an individual's test performance, like all cognitive tests, IQ is never an exact measure, but will always have a degree of inbuilt error variance.

At the next level is a series of specific cognitive domains, including memory, language, visual and executive functions; all except the last are familiar terms. Although there is less than total agreement amongst psychologists, the label of "executive functions" is generally used to refer to a set of abilities that enable people to plan, initiate and carry through goal-directed behavior in organized and "thought out" ways. Psychologists also draw a distinction between high-level cognitive functions such as integration, synthesis, planning, and organizing, and more basic, "low-level" cognitive functions such as processing information from the senses. These cognitive domains are not mutually exclusive (memories are often stored in linguistic forms, reasoning may involve visual stimuli) but they tend to be assessed using different tests, so it is convenient to label them separately. There are some mental processes like attention regulation and the maintenance of goal direction that operate across a number of other cognitive domains.

Each cognitive domain, in turn, can be decomposed into a series of components or sub-processes that are important in determining the level of function within that domain. Some of these refer to different processes, such as the understanding versus the expression of language or recalling information versus recognizing it. Others refer to the information on which the process acts, auditory versus visual stimuli, for example. In order to fully characterize performance within a cognitive domain it is necessary to administer a series of tests designed to evaluate key components. One should be suspicious of studies that claim to have assessed a complex cognitive domain, and then only report scores on a single test. If, however, an hypothesis predicted a very selective effect on the function this test assesses, this might be acceptable; even then, however, one might want to include another type of test from this domain to determine whether the hypothesized function was affected selectively. We will describe both batteries of tests that assess a variety of aspects of a cognitive domain as well as some commonly used single tests. Finally, there are very specific tests that are designed to assess basic components of cognition, often across domains, and usually developed in a research context in line with a specific hypothesis or model.

As a general point, it is important to note again that when we assess cognition in a child we are measuring behavior, i.e., performance, rather than competence. Children are born with intellectual potential that needs a stimulating environment to be fully realized. The quality of nutrition that the child receives will be a contributing factor to the quality of the environment. Hebb [19] used the terms Intelligence A and Intelligence B to distinguish between potential and the measured level of cognitive development. Intelligence A refers to the innate potential for intellectual development and can never be directly observed. What we observe, behavior on an intelligence test, for example, is Intelligence B, reflecting the level of development at the time of testing. A simple example: two individuals with the same Intelligence A are tested on vocabulary, a component of most intelligence tests. If a word has never been heard in the environment of one but is common in the other, the two will differ in the behavior they display on the test (Intelligence B), although their level of Intelligence A is the same. Intelligence A and B are not independent of each other, of course, since Intelligence A will contribute to the development of Intelligence B but they are not the same and it is important to bear this in mind. We have no way, at present, of measuring potential, although one can see from the above example that measuring nonverbal performance in children thought to have had an impoverished cultural environment, in which language experience was sparse, may give a better estimate of their overall ability. Although we have been discussing intelligence, the principle that we are assessing behavior, and not potential, applies to all aspects of cognition.

These cognitive domains exist across the age spectrum, and three papers published in 2005 [4, 22, 33] relevant to nutritionists, contain an interesting discussion about these same issues in adults. Children may need rather different methods of assessment, however, particularly if they are pre-verbal. One difference between the cognitive structure of adults and children is that the neural processes that underlie cognition are thought to be more focally represented in adults [14]. Although there is much disagreement about the degree of such representation in infancy, there is general consensus that, as development continues, cognitive representation becomes more modular. The practical result of this is that an intervention such as nutrition that affects a specific neural area may have different consequences in children at different ages and in adults. Cognitive domains, like brain regions, follow their own trajectories of development so that adult levels of performance are achieved at different chronological ages.

Test choice in relation to study design

The topic of designing a study is a complex one and, while not the focus of the present paper, the choice of tests to be used in a study is not independent of its design, so we will mention several relevant issues here. Crucially, researchers must be clear whether they are interested in global nutritional effects across a range of cognitive functions, or whether their interest is focused on a specific set or subset within this wide spectrum. If the interest is in global effects, then it is unlikely that a global measure such as an IQ result will have sufficient power to determine any but the most potent effects, unless very large sample sizes are feasible and controls for factors such as socioeconomic background can be implemented.

If the interest is in specific cognitive functions, and in general such more tightly hypothesis-driven studies are appropriate in this field, then the researcher should be careful to ensure that the psychometric parameters of the candidate test are adequate. What this means is not just that the test shows adequate reliability and discrimination power, but that it also has been demonstrated to be a valid measure of the function that it purports to assess. Such information is available from the publishers of tests or from test author journal publications. Tests without a robust evidential psychometric base should be viewed with caution. We discuss this further below.

Historically, the most common cognitive outcome measure in nutrition studies, at least those conducted in a medical context, has been some global measure, like IQ. Studies are starting to include more specific measures, however, with the realization that a global measure by itself can never adequately describe cognition and many more specific effects of nutrition may be missed by reliance on the IQ score or equivalent. Since there is not a large literature to consult in order to determine which cognitive domains to assess, a useful approach may be to include a battery that covers a variety of these (see below); if effects are shown in a particular domain, subsequent studies may attempt to further delineate the component processes that are involved. Sometimes, we have some knowledge of which neural areas are affected by the nutritional intervention and this can guide us to the choice of cognitive domains to investigate. If the frontal lobes are known to be involved, for example, we would predict that executive functions would be worth assessing. Unfortunately, more often than not, we do not have this information.

Where there are likely to be wide ranges of verbal abilities and cultural/ethnic backgrounds in a study sample population, it would be appropriate to consider using nonverbal performance tests to reduce additional variance arising from confounding sources, since language-based tests are much more vulnerable to such confounds.

Some studies are designed to measure performance at one time point, but others repeat the measurement, for example, before and after some period of intervention. This type of test–retest design is often relevant to studying nutritional effects. If the test is to be used on more than one occasion, it is worth considering tests that have two or more equivalent forms, thus helping to minimize practice effects. If no such test is appropriate for the study, then the initial score can be used as a covariate in the statistical analysis of the data. The effect of this is to look at the second test scores while statistically adjusting the first scores so that these are equal across subjects. Testing performance on more than one occasion allows the determination of the trajectory of development of a function. An interesting research question is the effect of nutritional intervention on cognitive developmental trajectories in terms of characteristics like their shape and slope.

There is one type of design that is particularly appropriate to nutritional studies; it is mentioned here because it is considered to be the "gold standard" for research in this field. This design is the double-blind randomized, controlled trial, usually referred to as an RCT [1]. It is rooted in the pharmacological model and is probably less widely used in other areas of psychological research. Samples of individuals who are eligible for inclusion in the study are randomly assigned, using some coded method, to a treatment or a control group, i.e., they receive the supplement or a placebo. Because of the random nature of the assignment, the assumption is made that many extraneous factors that might affect outcome but are not directly measured (for example, parental IQ, regular diet) will be the same between the two groups; any differences between the groups will be "washed out" by the randomization. Double-blind refers to the fact that neither the subjects in the experiment nor the experimenters themselves are aware of which subjects have been assigned to which group. The code revealing this is broken once the study is completed. In some studies, the treatment group is compared to a group that has been given no treatment rather than a placebo. In such cases, assignment to groups may still be random, but the subjects know which to which arm of the trial they have been assigned. It is important that this information is not known to the tester since even the most professional test administrator can be subtly biased in their assessment by knowing to which arm of a trial the subject has been assigned.

Choosing tests

As noted, we have focused on tests that could be used within the context of children living in developed countries and consuming a typical Western diet. The language in which tests are available, overwhelmingly English, is also a factor to consider in a European context. It is not possible for a researcher to simply translate a test from one language into another and assume that they are equivalent [17]. Even a simple task like translating a list of single words for presentation in a memory task may result in important differences, because words vary in their frequency of usage and their semantic associations across languages. The effect is that the individuals given the translated task are, in practice, being given a different test. This may make comparisons with studies using the test in its "native language" difficult and, at worst, may lead to misleading results. A certain amount of discretion must be used. Substituting a UK English word for an American English word—"biscuit" for "cookie" or "flat" for "apartment"—will probably not result in differences in performance. The more two languages and cultures diverge, however, the more caution must be exercised. It is easier when the tests are nonverbal in nature but, even then, equivalence may not be achieved. Visual material, for example, may picture objects that are unfamiliar in a different cultural context. The experimenter should always be alert to the possible influence of test materials on outcome measures. If, as is usual in nutrition studies, the goal of the study is to compare groups on outcome only, without relating their scores to the population norms, then it is permissible to use translated items, provided that this is made clear when the research is reported. The crucial thing is that all participants are given the same test. Nevertheless, the researcher must be alive to the possibility that changing items by translation may affect the validity of the test. We include a table of some common tests and the European languages in which psychometrically adequate versions are available (Appendix 1). Many American tests have Spanish versions because of the size of the Hispanic population. Our choice of tests was guided by the two principles of (a) use in a Western European context and (b) language availability, in conjunction with an attempt to cover the main cognitive domains.

Another important consideration is whether a test is standardized or not. Standardized tests have norms that reflect the scores that the population for which it is intended achieve on the test; the average performance for the group is usually, but not always, assigned a score of 10 or 100. The larger and more representative of the population the better; tests with norms derived from small groups may not be so reliable in reflecting the performance of an individual relative to the group. The use of standardized tests allows one to assess where an individual child's performance lies with respect to the population but if this is not a concern of the study, then nonstandardized tests may be suitable. This is usually the case when two groups, such as one receiving a nutritional supplement and another in a placebo condition, are compared and our interest is in determining whether they differ significantly on some cognitive outcome but not in relating this to a broader population. Providing that they are used in the proper context, nonstandardized tests should certainly be considered, although it will still be crucial to determine that the psychometric properties are adequate. Another advantage of standardized tests, relevant in some studies, is that they allow comparison between scores on different tests. It is not obvious that a child receiving raw scores of 38/62 on an arithmetic test and 21/40 on a test of spelling, for example, is at the same level on the two tests of attainment. Standardized tests, however, produce standard scores and if we know that each of these raw scores has the same standard score, then we can legitimately draw this conclusion.

Researchers should check to see how recently a test has been normed. Norms produced many years previous may give less accurate estimates of current population performance levels. The Flynn effect [16] refers to a rise in average IQ scores in a population that occurs over time. The average rate of this rise seems to be around 3 IQ points every decade. There has been much discussion about this phenomenon and about the possible causes, one of which is suggested to be improved nutrition. A practical effect is that norms become outdated and the "average" increases with time, i.e., 100 will no longer be the mean IQ score in a population. Test manufacturers, therefore, produce new norms from time to time and the use of outdated tests my give misleading measurements. Although only demonstrated for IQ tests, some people assume that other tests of cognition show increases over time as well.

To be of scientific value, psychological tests need to meet a set of "psychometric" requirements [28]. In order to allow us to draw valid conclusions, measures must be both valid and reliable. A test's validity reflects the degree to which it actually measures what it purports to do. There are various types of validity—for example, content, construct, predictive—but all are concerned with determining how confident we may be in using the test results to make inferences. Reliability refers to the consistency of test measurements—would a child taking the same test twice receive the same score on both occasions? We can think of this in terms of the stability of the results, expressed in terms of a reliability coefficient. As a general rule, we would expect the reliability coefficient of a cognitive test to be at least 0.80. Standardized tests will have information about these indices in the technical manuals that accompany the tests. The manual for the WISC-III, for example, reports reliability coefficients of 0.95, 0.91 and 0.96 for VIQ, PIQ and FSIQ, respectively. It also shows that the reliability coefficients for the subtests are generally lower than for the global measures, illustrating that greater confidence may be placed on the accuracy of the IQ score than for that of an individual subtest. The standard error of measurement of a test and the confidence intervals surrounding a score are also useful indicators of a test's reliability. Such information will not usually be available for nonstandardized tests and the researcher will have to use more indirect sources: do the test scores correlate adequately with standardized measures of the same function? Do the test results covary with results from other tests in an expected way?

Tests must also be sensitive so that they are able to measure what we want to observe. Many of the tests described below have been developed in a neuropsychological context, designed to detect changes in cognitive behavior in individuals with neurological disorders or who have suffered brain trauma, but they may not be so useful when the changes in cognitive behavior are more subtle, as in many nutritional studies. There are no standardized tests that have been specifically designed for use in nutrition studies but several of the listed tests have been developed to be particularly sensitive to small changes in performance and have been used successfully in drug trials with normal volunteers. This may make them particularly suitable for nutrition research. A good general guide is to consult the literature to find tests that have proved sensitive in similar situations in the past. A useful reference is Westenhoefer et al. [36] that contains a description of tests that have been used in previous nutritional studies. If such information is not available, it is worth conducting a pilot study first to see if the proposed measures are sensitive enough for the study.

Tests for school-age children are usually in either pencil-and-paper form or presented by computer. The variety of the former is greater and their administration does not depend on costly equipment or proximity to a computer, so they are generally more flexible. Both types of test vary in the degree of experimenter involvement, ranging from tests that are completely automated to those in which each item is presented by the researcher. The latter must be administered individually but some tests can be given to groups. Computer presentation is generally more objective and finer measurements of variables like reaction time are possible to obtain, but it may not allow for detailed observation of test behavior. Priorities must be set by the researcher.

Whatever type of test is used, it is imperative that standard instructions are used with each child. In a sense, each test administration is akin to a psychological "experiment" in which a child is placed in a standard situation and fixed materials are used to elicit samples of behavior. The instructions given to the child are a component of this standard situation and, unless the same instructions are used *verbatim* with each child, the results between children will not be comparable. Nonpsychologists do not always realize this and, for the best of motives, try to aid children to do their best by "helping them along" with additional instructions or examples. Tests for children under two years of age, before language acquisition has occurred, are rather different in form; by necessity, most such tests consist mainly of "performance" items. Such items involve compliance with the tester's requests and are, therefore, susceptible to tester bias. To avoid this and to acquire the skills necessary for smooth administration of the test, extensive training is usually required.

The main factors in determining the choice of tests will be the age of the participants, the outcome to be measured and the language of administration. In practice, the choice of tests is also often constrained by the time available for testing.

Cognitive tests

We present below a selection of tests, organized by levels of assessment, from the wide variety available. Where possible, we have selected tests that are available in European languages as well as in English (see Appendix 1 for a list of common tests available in a variety of European languages). We have included a list of tests described in the text in Appendix 2, providing information about their age range and time to administer, as well as a list of websites for test publishers. The publisher of a specific test is usually easily obtained by using an internet search engine. Publishers may be approached to find out whether a test of interest is available locally and/or in the local language. Some tests will be available in different languages but will have normative data only for the original population; decisions about their suitability must reflect the points we have described above.

Level I—overall ability

Tests of overall ability attempt to describe the general level of intellectual function that a child displays. It is assumed that this represents some sort of intellectual reserve that is deployed across more specific cognitive domains. Tests that attempt to produce an overall assessment of cognitive ability reflect this as they are usually, although not always, composed of a variety of subtests that have been found to have shared variance.

Measuring overall ability forms an important part of most test batteries for several reasons. These are the only measures that some nutrition studies report, making them necessary for comparison with the literature. They also provide a standard against which specific tests can be evaluated. How we interpret a score on a reading test of 80 in a child with an IQ of 100 is likely to be different in a child with the same reading score but an IQ of 80. We also often need to rule out overall intellectual differences between groups as an explanation for more specific differences in cognition that we wish to attribute to nutritional intervention. Overall tests, however, also have important limitations. When they are composed of various subtests, the same overall IQ score may reflect different patterns of performance, sometimes quite substantial, as different combinations of subtest scores may add up to the same total score. It is also wrong to assume that two groups of children with the same overall score will necessarily perform at the same level on specific tests—identical IQs may mask different patterns of ability on other tests. There are some cognitive domains, particularly executive functions, which are under-represented on most intelligence tests. As long as the limitations are recognized, however, a measure of overall ability can be a useful basis for a cognitive assessment. There is some evidence that early variations in diet may selectively influence VIQ scores [21, 36] while later interventions may have more effect on PIQ [3].

Birth—2.5 years

In children of this age, tests of general ability are usually thought to assess the level of neurodevelopment rather than intelligence. Studies of nutritional effects in this young population are especially significant since the plasticity of neural development in this period renders the young child particularly sensitive to nutritional factors. Probably the most widely used test in the English-speaking world is the Bayley Scales of Infant and Toddler Development (Bayley-III: age range: 1–42 months; Test 1). The test covers three major areas of development using a high degree of child–tester interaction: cognitive; language (receptive and expressive) and motor (fine and gross). In addition, two scales use parent questionnaires: social-emotional and adaptive behavior. The Behavior rating scale is an optional sixth subtest. As well as composite scores, the use of subtest scaled scores allows discrepancy analysis to determine areas of strength and weakness. Test administration is, of course, individual. Former versions of the test provided two composite scores: the mental development index and the psychomotor development index. Although the number of languages in which it is available is limited, the performance nature of many of the items means that many could be used without translation, although the level of a child's familiarity with specific test materials may still influence performance. Learning to administer tests like the Bayley-III to infants and young children involves a high-degree of training, necessary if any confidence is to be placed in the test results.

Scale 1 of The Griffiths Mental Development Scales (Test 2) covers the 0 to 2-years age range and provides measures in five areas of cognitive development: locomotor, personal-social, hearing and speech, eye and hand coordination and performance. It is a UK based test and is widely used by medical researchers and practitioners but is not available in other languages. A new American test, the Mullen (Test 3) is starting to be used widely by researchers. It can be used from birth to 68 months of age. It measures fine and gross motor function, visual reception, expressive and receptive language, but is only available at present in English.

There is one more type of assessment that is often used as a measure of general ability in infants because, although limited in scope to one type of function, there is some evidence that it may serve as an indicator of intelligence scores later in childhood (although reported correlations vary and are mostly low [6]). This assessment measures infants' visual novelty preference or visual recognition memory. The researcher exposes the infant to a target visual stimulus until the infant habituates to it and looks away. The target stimulus is presented again along with a novel stimulus and the relative proportion of time spent looking at the two stimuli is measured. It has been argued that infants who tend to habituate more quickly and spend more time looking at the novel stimulus rather than the familiar stimulus also tend to perform well on tests of general intelligence at later ages. It is hypothesized that the processes underlying these early visual novelty preferences are basic to intellectual performance and are the same as those that will underlie intelligent behavior later in the lifespan. One such formalized procedure is the Fagan Test of Infant Intelligence (FTII, Test 4).

Over 2.5 years

The choice of tests to assess intelligence in children older than 2.5 years is much wider. Most are composed of a battery of subtests, sampling a broad spectrum of behavior, the scores of which are then combined to give a single measurement, the full scale IQ (FSIQ). These subtests can usually be grouped into those that involve verbal material and those in which the stimuli are nonverbal—most tests provide separate measures of the child's ability in these two broad domains as well as FSIQ. By far the most widely used tests worldwide, available in the greatest variety of European languages (see Appendix 1), are the Wechsler scales. Two of these cover the age range from 2 years 6 months to 16 years 11 months: the Wechsler Preschool and Primary Scale of Intelligence (WPPSI-III, Test 5) and the Wechsler intelligence Scale for Children (WISC-IV, Test 6) for the older group. An adult test (the Wechsler Adult Intelligence Scale—WAIS) allows assessment over the whole age range, making the Wechsler tests suitable for longitudinal studies. All Wechsler tests provide a verbal IQ (VIQ) and a performance IQ (PIQ) in addition to the FSIQ. The Third edition of the WPPSI is somewhat different in format to its predecessors, in that it has two different forms to be used with the 2 years 6 months–3 years 11 months and the 4 years 0 months–7 years 3 months age groups. In addition to IQ scores and a general language composite score for both groups, a processing speed quotient can also be obtained for the older group only. The WISC-IV, in addition to IQ scores, also provides four index scores allowing a finer-grained analysis of cognitive ability: verbal comprehension, perceptual reasoning, working memory and processing speed. The tests allow score patterns across subtests to be analyzed. If time constraints exist, since they take around 90 min to administer in full, the Wechsler Abbreviated Scale of Intelligence (WASI) which provides VIQ, PIQ and FSIQ scores based on four subtests may prove useful (Test 7). Other widely used intelligence tests in English are: the Kaufman Assessment Battery for Children (KABC-II, Test 8), the British Ability Scales (BAS2, Test 9) and the Differential Ability Scales (DAS, Test 10), an American test based on the BAS; all are as time-intensive as the Wechsler scales.

All the above tests share a similar structure, but there are other options worth noting. The Cognitive Abilities Test (CAT3, Test 11) is designed for group administration in school-age children. It measures reasoning in three domains: verbal, nonverbal and numerical, and could prove useful in studies where some measure of overall ability is needed but individual face-to-face testing is not feasible. The Leiter International Performance Scale (Leiter-R, Test 12) is entirely nonverbal and performance is, therefore, independent of the language of administration. It uses visualization and reasoning scales to measure IQ and also has scales that measure attention and memory. Finally, Raven's Progressive Matrices (RPM, Test 13) is a widely used nonverbal test of intelligence which was designed as a measure of "g". A colored form for use with children and a standard form for older participants are available; in both, the oral instructions are very simple. The test does not measure verbal ability although vocabulary scales are available to use with it, but only in English. The correlations between scores on the RPM and the Wechsler tests tend to be in the 0.70–0.80 range. The test can be administered to a group and it is available with manuals in a wide range of European languages (see Appendix 1).

Another useful nonverbal intelligence test is the Snijders–Oomen (SON-R, Test 14). Seven subtests measure four areas of cognition: abstract reasoning (categories and analogies), concrete reasoning (situations and stories), spatial abilities (mosaics and patterns) and perceptual competence (hidden pictures). Although the time needed to administer the SON-R is, on average, 1.5 h, there is a shortened version consisting of four subtests: categories, mosaics, situations and analogies, taking around 45 min.

The most recently published nonverbal test is the Wechsler Nonverbal Scale of Ability (WNV, Test 15), which gives an overall IQ. It has two forms, depending on age at assessment. At each age-level, there are two-subtest (administration time 10–20 min) and four-subtest (30–45 min) batteries. It has been designed for use in different countries and it is claimed that the native language of the child is irrelevant to performance. All directions are given in pictorial form so there is no need to adapt the administration of the test or to pantomime instructions. Initially, norms are available for USA and Canada only, but it could be used for group comparisons in many countries.

Level II—batteries measuring multiple cognitive domains

We are concerned here with obtaining measures of performance on a variety of cognitive domains, usually with a series of sub-tests within each domain. The main advantage to this approach is that norms for all the sub-tests have been obtained from the same population and they can thus be compared easily and reliably with each other. In nutrition studies, we might hypothesize that one particular domain would be affected more than others and the use of such a battery would allow us to confirm the selectivity of the intervention. There are no such batteries for use with children under 2.5 years of age since the functions that these tests measure are not considered directly testable in infants and toddlers. Even in older children, the choice is much narrower than for tests of overall ability.

Over 3 years

The best known is the NEPSY-II (from NEuroPSYchology), originally designed to measure the neuropsychological status of children aged 3–12 years (Test 16). It covers five domains: attention/executive function, language, memory and learning, sensorimotor function and visuospatial processing. A series of core subtests for each domain measures key component processes, as well as supplementary tests if more detailed assessment is required. For example, the memory and learning subtests are: immediate memory for sentences, immediate and delayed recall of names and faces, list learning and narrative memory under free recall and cued conditions. Five domain scores can be calculated, provided that the core subtests have been administered, but there is no overall score as there is in the case of IQ, since it is the pattern of performance that is of interest. A new sixth domain, social perception, has been added in the second edition and the upper age limit has been extended to 16 years. The test is published in the USA, although it was developed originally in Finland and it is available in some European languages with local norms.

Another option is the CANTAB (CAmbridge Neuropsychological Test Automated Battery), a computer-based battery of tests originally developed for adults but now with norms for children (Test 17). Until recently, all test stimuli were nonverbal and described as independent of language and culturally nonbiased; several verbal sub-tests have been added recently. These 22 subtests are grouped into batteries to provide measures of domains like executive function; it is possible to create customized batteries for any particular study. A unique feature of the CANTAB is that performance on the various subtests is known to correlate with activity in particular regions of the brain, one reason why the CANTAB is very widely used in academic research settings. All responses are automatically recorded during performance, with precise timing where relevant, and a wide variety of measures is provided for each subtest. The CANTAB has proved to be sensitive to small changes in behavior, such as those associated with medication in normal subjects and this feature has also made it a popular choice for research studies and a good candidate for use in nutrition studies. While the new verbal sub-tests may not be suitable for use with non English participants, the nonverbal tests should be.

Another UK-based computer battery is Cognitive Drug Research, known as CDR (Test 18). As the name implies, this battery was developed for use in drug trials, which may make it suitable for nutrition studies. It has been used very little in academic research, since the company is usually contracted to conduct the studies itself. The company is flexible in its approach, however, and attempts to meet researchers' needs. The battery was developed for use with adults but most subtests would probably be suitable for older children and possibly with younger ones after a pilot study. Similar in structure to CANTAB, the core subtests measure motor control, attention, working and secondary memory with a series of supplementary tasks to provide further information. Although the tests are mainly nonverbal, the few that use verbal stimuli are available in most European languages.

Level III—individual cognitive domains

We may have an hypothesis predicting that function in a particular cognitive domain will be affected by our nutritional intervention or even that some aspect within a domain will be selectively affected, leaving other aspects intact. In cases such as this, we will want to administer a test that measures a variety of cognitive components within a domain. Many studies, however, use just one or two tests, often combining individual tests from different domains to form an ad hoc battery. This means, of course, that the researcher has chosen to evaluate only one or two of the many component processes of the cognitive domain. This may be appropriate if there is an hypothesis predicting that certain very specific processes will be affected selectively by the treatment. The tests will not have been normed on the same population, however, making comparison between test scores more difficult than if a multiple-component test is used. It is usually best, under these circumstances, to also include a test of another function in the same domain *not* hypothesized to be affected, in order to establish the selectivity of the effect. Examples of both types of test, multiple and single component, are described in this section. Many more may be found by consulting test catalogues and textbooks. Few such tests are available for the younger age group.

It is worth noting that some Level II Batteries, notably the NEPSY, can be used to assess single domains in isolation and single tests from these domains can also be used. Some of the tests are suitable for 3 to 4-year-olds and these are often the only standardized tests available for this age group. Almost all the subtests can be used between 5 and 12 years, so it is worth considering such batteries as a source of useful tests. They must, however, allow performance on individual tests and domains to be quantified and not have a single score as the only outcome measure.

Memory

Memory refers to the registration, storage, retention and later retrieval of information to which the person has been exposed and is a complex system made up of components that may be dissociated from one another [2]. Thus, a comprehensive assessment of memory will need to have a series of sub-tests. We may, for example, measure memory for verbal versus nonverbal material, vary the presentation so that some stimuli are visual and some auditory, test different modes of retrieval such as recall or recognition, and vary the interval between first presentation of the material and subsequent recall.

Multiple component tests

No such comprehensive memory batteries are available for children under five.

Age 5 and over There are three widely used batteries to measure memory function: childrens' Memory Scale (CMS, Test 19), test Of Memory And Learning (TOMAL, Test 20) and the wide Ranging Assessment of Memory and Learning (WRAML-2, Test 21). Although there are minor differences between them, we will describe the CMS to give some idea of how they are all structured. The CMS uses visual and verbal material in short and long-delay conditions. Recall, recognition and working memory are assessed and learning and attentional characteristics are described. A core of six subtests can be administered with additional subtests for more complete evaluation. A series of composite index scores (for example, verbal immediate, verbal delayed, attention/concentration) is produced, as well as a general memory index. There are tables that allow the determination of whether dissociations between different aspects of memory are statistically meaningful, for example, attentional factors versus memory per se. A useful feature of the CMS is that it allows memory scores to be linked to performance on the Wechsler scales so that we can see if memory function is in line with expectations from IQ.

Different in structure and purpose to the above is The Rivermead Behavioral Memory Test (RBMT-II; RBMT-III due in 2008, Test 22), designed to assess memory skills related to everyday situations. It claims to be more ecologically valid and includes subtests measuring the recognition of faces and objects, "prospective" memory for future activities, recall of a route and a story, both immediately and after a delay and items measuring orientation in time and place. Quick to administer, this test provides two measures: a screening score that offers a simple way to estimate the extent of everyday memory problems, and a profile score which gives a more sensitive measure. Four parallel forms of the test are included, allowing accurate measures of change over time in test–retest designs. The adult version has been used with adolescents from 11 to 15 years of age. A children's version (RBMT-C; Test 22) is available for ages 5–10.

Single component tests

Under 3 years Memory in children under 3 years of age is rarely assessed in the same way that it is in older children and the measures are often primarily designed for other purposes. The Bus Story (Test 23), for example, involves telling the child a story while looking at a series of accompanying pictures and then asking the child to tell the story back using the pictures as prompts; this was devised as a language test, although it obviously has a large memory component. Tests of visual recognition memory such as that of Fagan were mentioned above as measures of overall cognitive ability. Tests like the BSID and the WPPSI contain some items that measure memory. Uncertainty about the processes involved in memory development in young children have hampered the development of appropriate standardized tests although research paradigms may be of interest (see below).

Over 5 years The Children's Auditory Verbal Learning Test (CAVLT-2, Test 24) measures auditory verbal learning and memory abilities in children and adolescents aged 7–18 years, using a word-list recall paradigm. Although it is a single test, it uses one recognition and two free-recall memory word lists and yields measures of immediate memory span, level of learning, immediate recall, delayed recall, recognition accuracy, and total intrusions.

The Visual Aural Digit Span test (VADS, Test 25) is suitable for school-aged children and consists of four subtests in which numeric sequences involving the digits 1–9 are used as stimuli. The child orally repeats or writes the sequences from memory. The first subtest requires oral repetition of orally presented digits. In the second subtest, the child must orally repeat digits that are presented visually. The third subtest requires the child to write digits that are orally presented while the fourth subtest involves writing digits that are visually presented. This provides a great deal of information (11 measures) compared to the usual digit span test involving oral presentation and recall only, although the latter has the advantage of recall both in order of presentation of the stimuli and backwards and can result in some interesting dissociations between performance in the two conditions. The VADS is easily adapted for use in different languages.

The Rey–Osterrieth (RCFT, Test 26) test uses complex visual material to measure visuospatial ability and visuospatial memory. It is suitable for children from 6 years as well as for adults. Although there are various research versions of the test and a variety of scoring systems, it is now available as a commercial test with appropriate norms. The figure is copied when first presented and the child is then required to reproduce it from memory both immediately and after a 30 min delay. There is also a recognition trial in which the child has to select which elements of the figure have been seen before by choosing between correct items and incorrect foils. The RCFT measures visuospatial construction ability (copying trial) and visuospatial memory (immediate and delayed memory, recognition). Also available is the developmental scoring system for the Rey–Osterrieth Complex Figure (DSS-ROCF) which is suitable for children aged 5–14 years. It is specifically concerned with evaluating performance within a developmental context and emphasizes qualitative aspects of the child's drawing such as organization and style. Determining the maturity of the child's performance may be of interest in some studies.

Language

Language and verbal processing play a key role, of course, in the assessment of VIQ. Such tests, however, do not assess language per se but rather how it reflects verbal knowledge and verbal reasoning. A child asked to define a word on an IQ vocabulary sub-test, for example, may use ungrammatical language but is credited for the item if the reply contains the correct information. Language assessment, on the other hand, concentrates more on the formal and pragmatic aspects of verbal usage. Given the obvious impact of past learning experiences, ethnic background, and other factors on language abilities, language-based tests provide highly confounded measures of cognitive abilities. Because of the importance of language in educational achievement and its use in verbal reasoning and verbal problem-solving, however, it is important to determine whether nutrition affects this domain of cognitive function. Any assessment incorporating verbal measures must ensure that adequate controls are implemented to minimize or account for confounds; details of maternal and paternal education and achievement, socio-economic status and information about the level of stimulation in the home environment (preferably using a measure such as the HOME inventory [13] should be collected and used in analyzes of the data). Because of the sensitivity to cultural differences, it is best to use a test devised for use in a particular country. Speech and language professionals should be able to provide advice for suitable materials available locally.

Multiple component tests

At the very least, a comprehensive language assessment should provide measures of both receptive and expressive language. It may also measure components within each of these two domains.

Under 3 years The MacArthur–Bates Communicative Development Inventories (Test 27) are checklists reflecting language comprehension and expression that are completed by parents; scores obtained from them have proved to correlate highly with assessments by language professionals. The CDI: words and gestures (infant form) and CDI: words and sentences (toddler form) are used from 8 to 30 months. There is an upward extension, the CDI-III, suitable for assessing language skills in children aged 30–37 months. Typically, vocabulary checklists are presented and the parent indicates which items the child comprehends and produces. For older children, there are questions about the use of gestures, combining words, the degree of grammatical complexity in the child's speech, and yes/no questions concerning semantics, pragmatics, and comprehension. The tests have been very widely used and there is a large literature available. Relevant to this article is the test availability in a wide variety of European languages. As well, the website contains useful information about test development in different languages which has relevance beyond this particular test.

The NEPSY language domain includes four subtests assessing receptive and expressive language for 3 and 4-year-olds.

Over 3 years The Clinical Evaluation of Language Fundamentals (CELF 4th edition, Test 28) is an example of a test designed to measure a wide variety of language components, containing both core and supplementary subtests. The four core subtests provide a total language score as well as composite scores for language structure, language content, language content and memory and working memory. Additional tests measure the skills and behaviors that underlie the child's language: phonological awareness, rapid automatic naming, digit span, sequences, word associations and memory. These are useful in trying to determine whether there is a selective effect of, for example, nutrition on certain aspects of overall language function. There is also a preschool version (CELF Preschool, Test 28). It too measures a wide range of language components.

Single component tests

Over 3 years Probably the single most widely used test to assess language is an American test, the Peabody Picture Vocabulary Test; its incarnation in the UK is called the British Picture Vocabulary Scale (BPVS-II, Test 29). This is a quick measure of receptive language that estimates the size of the child's lexicon. Pictures are presented, four to a page, and the child must point to the one that matches a spoken word; no reading, writing or speaking are required of the child. The stimulus items range from concrete to abstract. It does not measure expressive language but is often used to estimate overall verbal ability (the highest correlation between a subtest and overall IQ is that for vocabulary).

The Test for Reception Of Grammar (TROG-2, Test 30) tests the understanding of 20 English grammatical concepts, four times each, using a picture multiple-choice format. The child points to the picture that matches a spoken phrase or sentence by the tester so, like the BPVS-II, this is a measure of receptive language but at a more complex level than single-word knowledge.

The most usual way to measure expressive language, using a single test, is to administer a naming test. The child is presented with a series of pictures of objects, usually sequenced in order of difficulty, and has to produce the correct label for the object. There is no one test that is dominant in this field but the Word Finding Vocabulary Test (Test 31) is typical. More complex expressive language in children aged 3–8 years is often assessed by the Bus Story Test (Test 23). As described above, the child is told a story while looking at accompanying pictures and then re-tells the story. The age level of consecutive speech used in this re-telling can be assessed from the information content, sentence length and grammatical usage. Many language tests can be found by consulting educational suppliers such as Taskmaster (see websites list).

Attention

The term "attention" refers to our ability to select and focus on specific aspects of the stimuli, both internal and external, to which we are exposed. It is closely allied to working memory, the system that allows us to operate cognitively on information held in consciousness. Attention, like other cognitive domains, is complex and current research indicates that different functions within this domain are underpinned by different neural systems. Commonly identified components of attention are focusing, sustaining and shifting [26] or, referring to similar concepts but using different names, detecting, alerting and orienting [30]. Tests should assess attention in both visual and auditory sensory modalities. There is increasing research interest in attention processes, in part because they are implicated in the performance of many higher level cognitive functions, but also because deficits and difficulties in attention regulation are prevalent in children and young people, with serious effects on learning and associated with behavioral disorders.

Disturbed attention is frequently associated with nutritional disorders.

Multiple component tests

There are no attention batteries for use with children under 6 years of age at present.

Over 6 years The Test of Everyday Attention for Children (TEA-Ch, Test 32) assesses a range of important attentional capacities and comprises nine subtests which measure children's abilities to attend selectively, to sustain attention, to divide attention between tasks, to switch attention and to inhibit verbal and motor responses. Both auditory and visual stimuli are used. The TEA-Ch is suitable for children aged 6–16 years, with separate norms for boys and girls.

Single component tests

Three to four years There are two NEPSY subtests that assess attention/executive function in this age group.

Over 4 years Conners' Continuous Performance Test (CPT II, Version 5, Test 33) is an attention test that is widely used in ADHD research and clinical assessments for people aged 6 years or older, and is quick to administer. Responses to target letters on a computer screen provide information about the specific type of deficits that may be present. For example, some response patterns suggest inattentiveness or impulsivity, while others may indicate activation/arousal problems or difficulties maintaining vigilance. This test can only be presented by computer. The Kiddie version (K-CPT, Test 33), for 4 to 5-year-olds, uses the same basic paradigm but with pictures of objects rather than words. It takes 7 min to administer. A similar test, also presented by computer, is the Test of Variables of Attention, for 4 to 80-year-olds, which has extensive norms for general, as well as clinical, populations (Test 34). One advantage of this test is that it has two versions, one visual (TOVA) and the other auditory (TOVA-A); it takes longer to administer (25–30 min) but provides more information. NonVerbal tests of attention can usually be used in any country.

The CANTAB, described above, includes several tests of attention that can be used alone and several subtests on the Wechsler IQ tests are also said to measure attention. The CMS, also described above, gives scores on an attention/concentration dimension.

Executive functions

We have earlier noted the significance of executive functions in cognitive performance. They may be thought of as higher-order cognitive skills that are associated with the child's ability to engage in independent, goal-directed behavior. A child with good executive function will demonstrate mental flexibility, will be able to form and maintain sets, to plan and self-monitor behavior, and will be able to inhibit impulsive responses, as well as demonstrating abstract reasoning, concept formation and rule learning. Deficits in executive function may show themselves as the lack of behavioral control (disinhibition), the inability to change behavior despite corrective feedback (perseveration), or inefficiency in problem-solving, difficulty in understanding cause–effect relationships and a lack of motivation. Although there are some experimental procedures for measuring problem-solving in young children, the full range of executive functions is thought to only develop in later childhood [10] and may continue well into adolescence as myelination of the frontal lobes proceeds. Although not tests, as such, two rating scales are available to evaluate a variety of executive functions in school and home settings in children aged 2–18 years. The Behavioral Rating of Executive Function-Preschool (BRIEF-P, Test 35) is for children aged 2–5 years 11 months and the BRIEF (Test 35) is for 5 to 18-year-olds. Scales are completed by parents, teachers or day care providers depending on the age of the child. They are quick to administer but provide a different type of information to that obtained from test administration.

Multiple component tests

Seven years and over The Behavioral Assessment of the Dysexecutive Syndrome in Children (BADS-C, Test 36) is a test battery that has recently become available to measure a range of executive functions in children aged 7–16 years. Its subtests assess the following functions:

- inflexibility and perseveration;
- novel problem solving;
- impulsivity;
- planning;
- the ability to utilize feedback and moderate one's behavior accordingly.

The subtests and materials are novel and of interest to most children. Each of the six subtests yields an age-scaled score, combined to form an overall scaled score which is classified into six categories, ranging from impaired to superior. There is also a 20 item questionnaire, the Dysexecutive Questionnaire for Children (DEX-C), to be completed by a parent and/or teacher who has frequent contact with the child. The subtests may be used separately but the real strength of the test lies in the assessment and profiling of different functions.

Single component tests

One of the most common ways to measure (one aspect of) executive function is to use a verbal fluency test. In this kind of test, the requirement is to produce as many different words as possible in a short time (usually 30–60 s), while following specified rules. Typically, the child is asked to produce words that begin with a certain letter, like M, (phonological condition) or belong to a certain category, like flowers (semantic condition), without repetition. This is considered to be a test of executive function because it involves a number of different cognitive processes that are disrupted in the dysexecutive syndrome. For example, the child must use a strategy to extract items from its own lexicon while inhibiting nonrelevant responses, and must not break the rules by using words from outside the category or repeating a work (perseveration). There are many versions of verbal fluency tests–a useful one, with norms for children, is one of the sub-tests of the NEPSY: the semantic condition is given to 3 to 12-year-olds, but the phonological condition is only used with children aged 7 years and above.

The Children's Color Trails Test (CCTT, Test 37) is suitable for 8 to 16-year-olds and assesses sustained attention, sequencing, and other executive functions. The child must connect stimuli scattered on a page by drawing lines between them in a specified sequence. It is suitable for use within any linguistic context as the stimuli are nonverbal and visual instructions may be used. It is quick to administer (5–7 min).

Another widely used single test is the Stroop color and Word Test (Test 38). The children's version is for 5 to 14-year-olds and takes 5 min to administer. It taps into the executive function processes of cognitive flexibility and resistance to interference from outside stimuli. The test consists of three basic parts: a word page consisting of the names of colors (red, green, blue) printed in black ink; a color page of semantically meaningless symbols (X) printed in different colored inks; a color–word page comprising the words from the first page, printed in the colors from the second page, with the restriction that the word and color do not match. The task is to look at each sheet, and move down the columns, reading words or naming the ink colors, as quickly as possible, within a given time limit. The test yields three scores, and Interference, based on the number of items completed.

Visual functions

This refers not to the assessment of vision itself, but rather to cognitive activities that involve the use of vision. For example, visuo-perceptual tasks measure how an individual perceives visual information, i.e., how the incoming visual stimuli information are interpreted by the brain. Visuo-motor tasks look at abilities that involve vision to guide and monitor motor behavior, while visuo-spatial tasks assess the ability to make correct judgments about spatial relations. There are relatively few tests to measure these abilities compared to memory and language, particularly in children. Some individual tests are widely used but there are few batteries. It is worth pointing out here, particularly if visual functions are the focus of research interest, that it can be quite difficult to isolate these functions from verbal processing. Even when presented with visual stimuli, humans tend to label them verbally and will often use verbal strategies when processing them.

Multiple component tests

Three years and over The Wide Range Assessment of Visual Motor Abilities (WRAVMA, Test 39) is designed for use between the ages of 3 and 17 years. It measures performance in three domains: visuo-motor (drawing), visual spatial (matching) and fine motor (peg moving). The latter is important since any motor difficulties may influence the score on visuo-motor tasks. It is quick to administer, taking about 10 min.

Single component tests

Two years and above The Beery–Buktenica Developmental Test of Visual–Motor Integration (Beery VMI, 5th edition, Test 40) assesses the extent to which individuals can integrate their visual and motor abilities. The short format and full format tests present drawings of geometric forms arranged in order of increasing difficulty that the individual is asked to copy. The age range is 2–18 years; the short format is often used with children aged 2–8 years. It contains two supplementary tests that help determine whether any deficit noted on the test is primarily due to difficulty with the visual or motor components. The initial test takes around 15 min, while the visual perception and motor coordination supplementary tests take 5 min each. It is linguistically neutral.

The Benton Judgment of Line Orientation Test (BJLO, Test 41) is a widely used neuropsychological test and is a "pure" measure of visuospatial judgment that is language-free. It assesses the accuracy of judgments of the slope of visually presented lines by requiring the testee to identify target stimuli consisting of pairs of lines of different slopes on a multiple choice display. Originally designed for adults, it can be used with children from 7 years upwards.

The NEPSY includes two measures of visuospatial function for use in 3 to 4-year-olds.

Level IV—research tests

These are tests that have been designed by researchers to fulfil the requirements of a particular study and are often theory-based. Referring to the scientific literature can prove very fruitful as a source of tests; it is usually possible to contact the test authors to gain additional information regarding psychometric properties, administration protocols and subsequent revisions, as well as permission to use the test material. Where comparisons between treatment groups are concerned, and reference to population norms is not necessary, such research tests may provide adequately sensitive indicators for very specific cognitive functions. Nevertheless, the researcher must not ignore the psychometric properties of such tests. Such tests are often particularly suitable for use with young persons since this age group is under-represented in the body of standardized tests. Many researchers, for example, use research paradigms to study function in domains like memory [24] and problem-solving [37] in toddlers.

The number of such tests is legion but to illustrate their character we will describe here just one—the inspection time paradigm, developed as a test of visual processing speed [27]. It is not concerned with how rapidly a response is made to the visual stimulus (reaction time) but rather with the amount of time that an individual needs to be exposed to the stimulus to make a correct response (inspection time). The stimulus is presented on a computer screen (Fig. 1) and the task is to press a key to indicate which of the two legs of the figure is longer, left or right. Given unlimited time, this is a cognitively trivial task but the duration of presentation ranges from 6 to 200 ms; at the lowest exposure times, performance is usually random. There are many repetitions of the figure at different stimulus

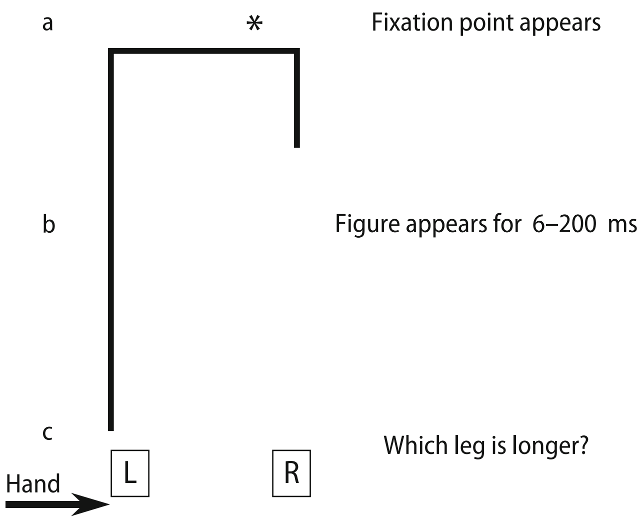

Fig. 1 In the inspection time task, the child fixates the star. The experimental figure then appears for a period of time ranging from 6 to 200 ms. The child presses either the *L* or *R* computer key to indicate which of the two legs of the figure is longer. A large number of trials with randomized duration of stimuli is presented

exposure times and the number of correct decisions at each duration is calculated. This sort of test is not standardized and the exact details of the presentation procedure may differ between studies, but it has led to a great deal of productive research.

Measuring indirect factors affecting cognition

As noted earlier, a multitude of factors can potentially affect children's cognitive performance when they are being assessed. This is especially true for younger children. The testing situation itself can be perceived in many different ways by children, and it should not be assumed that children understand that the aim of the tester is to evoke behavior that is as close as possible to the child's best achievable performance, that is, as close an approximation as possible to the child's underlying competence [30].

While most test manuals stress the importance of establishing 'rapport' with a child testee, the situation is inevitably one that places a child under stress, not least the stress of dealing with what have to be, by their very nature, difficult challenges, some of which the child will fail. The child is also face with a challenging social situation, where the motives of a strange and more powerful adult need to be assessed and related to in a way that the child thinks appropriate. Naturally, the child brings to such situations expectations of the most appropriate ways to behave, based on past experience. For some children, this may be to withdraw and be highly compliant; for others it may be to ingratiate themselves with the adult; "doing their best" is not a universal response. It will always be important to ensure that staff who are testing the participants in a study have adequate training in test administration with the relevant populations.

Motivation, temperament, illness, emotional disturbances, weather (windy conditions increase distractibility in young children), noise levels, lighting, time of day, time of year and many other factors can all impact on a child's performance. Apart from the first two, these factors are generally negative in their effects, degrading performance and hence leading to inaccuracy in measures of competence or potential. Because of these nonrandom effects, the noise that they introduce into the data does not simply reduce power, but may also produce erroneous and misleading apparent effects.

This brief discussion of such contextual effects is given here primarily to stress the importance of implementing controls or measuring potential confounds to minimize such sources of error, and to emphasize the need to use manualized procedures to ensure that protocols are consistently administered. More about such issues are to be found in books devoted to testing considerations [25, 33, 34].

Concluding remarks

The focus of this paper has been on offering practical advice with regard to selecting tests that can be used in nutrition studies, rather than on theoretical concepts. We have presented a variety of tests, using a levels-of-assessment approach to cognitive measurement, which may aid the researcher in selecting appropriate tests for their study.

We must emphasize that the more thought that goes into planning the study before data collection starts, the better the outcome will almost certainly be. This means articulating very specifically what the aim of the study is and choosing a design that will allow the question to be answered. It is a good idea to plan the way that the data will be analyzed ahead of time to ensure that the appropriate data are collected to allow the statistical analysis. Particular attention needs to be paid to trying to isolate factors that may confound the results. If there are socio-economic differences associated with the outcome of interest, for example, then that factor needs to be controlled, either by demonstrating that the groups do not differ or, if they do, by using it as a covariate in analysis. Either way, it is necessary to measure socio-economic status at the time of data collection and, as mentioned above, it is also useful to include some sort of measure like HOME [13].

If analyses of the data show that statistically significant differences exist between treatment and control groups, it is important to realize that the underlying effect sizes are of great importance in interpreting findings and publishing them. It can be misleading to report statistically significant but very small effect sizes in popular media or advertising since the public may not understand that other factors with much greater effect sizes may in practice overwhelm the reported effect. Although it may be found that a difference is statistically significant, it may nevertheless be so small in absolute size that it in practical, physiological and developmental terms it is of no significance whatever. In part this is because of the inherent instability of cognitive test measures over time. A mean difference of, say, 5 IQ points between two groups, for example, a treatment and a control group, is unlikely to be sustained as a group difference over time because of the normal fluctuations in individual IQ scores over time that can be expected. Longitudinal designs, where the persistence of effects over time can be assessed by repeated measures taken from the same individuals at successive time points, offer one solution to this issue.

For reasons discussed earlier, tests of cognitive functions can never be 100% reliable. The variance that even relatively reliable tests introduce into measurements always needs to be taken account of, and this is a consideration that will always push for larger and hence more expensive, sample sizes. At the same time, the issue of effect size is also relevant to decisions that have to be taken regarding the planned sample sizes for research studies. While it is now common practice to carry out statistical power calculations to estimate the minimum sample sizes required to show statistically significant results, this may lead to the use of excessively large and hence costly samples if reaching statistical significance is the only criterion. Smaller samples may be adequate to detect practically significant differences. Thorough examination of previously published results, or making use of a two-stage design where a pilot study seeks to indicate potential effect sizes, are recommended so that meaningful effect sizes can be planned for in the research design, rather than merely statistical significance.

The present is an exciting time for research into nutrition and cognition since brain imaging procedures allow us to examine whether cognitive effects may be found to have measurable neural substrates. Whether nutritional intervention can alter brain anatomy and/or physiology as well as behavior opens many new avenues of research. In tandem with this, new discoveries about the effects of gene × nutrient interactions on cognition also herald a new era of investigations. Such studies require collaboration between scientists trained in different disciplines to be fully effective. In all such studies, careful and appropriate measures of cognitive function are of prime importance. We hope some readers will go on to contribute to this body of literature.

Notes on References

In the reference section, we have included three books that are excellent sources of general information about tests [25–34]. All contain far more detail than space allows here. References to specific tests mentioned in the text are not presented in this section, but rather in Appendix 2 which summarizes some of the main characteristics of each. The tests are grouped by levels (this should help to identify areas where there are test "gaps") and appear in the order in which they do so in the text.

Acknowledgments The authors would like to thank the reviewers Maureen M. Black and Bonnie Kaplan for their useful comments and discussions. This work was commissioned by the Nutrition and Mental Performance Task Force of the European branch of the International Life Sciences Institute (ILSI Europe). Industry members of this task force are Barilla G. & R. Fratelli, Coca-Cola European Union Group, DSM, Groupe Danone, Kraft Foods, Nestlé, Südzucker/BENEO Group, Unilever and Wild Flavors. For further information about ILSI Europe, please call +32-2-771.00.14 or email: info@ilsieurope.be. The opinions expressed herein are those of the authors and do not necessarily represent the views of ILSI Europe.

Conflict of interests The authors have no financial or other interests that might conflict with the views expressed.

Appendix 1

Appendix 1 presents a list of widely used tests that are available in different European languages (all available in English) (Table 1).

Table 1 Table of test availability in different European languages for some widely used tests (all available in English)

Language	Tests available
Dutch	BSID-III, WPPSI-R, WISC-R, NEPSY
French	WISC-R, WPPSI-III
German	Raven's, WPPSI-III, WISC-IV
Danish	Raven's, BSID-III, WISC-III, WPPSI-R
Greek	WISC-III
Spanish	NEPSY, BSID-I, WISC-IV, Raven's
Hungarian	Raven's
Italian	Raven's, WPPSI-III, WISC-III

Appendix 2

Appendix 2 lists all tests described in the text in the order in which they appear in the text, i.e., grouped by levels. Summary information is given for each (Table 2).

Table 2 Table of tests referred to in the text, in the order in which they appear

Level of assessment	Test name	Age range	Time to administer	Notes
I—Overall ability				
1.	Bayley-III	1–42 months	30–90 min	Three administered scales: cognitive, language and motor; two questionnaires
2.	Griffiths Mental Development Scale	0–2 years	N/K	Measures for five domains of function: locomotor, personal-social, learning and speech, eye-hand coordination, performance
3.	Mullen Scales of Early Learning	Birth, 68 months	15 min at 1 year	Five scales: fine and gross motor function, visual reception, expressive and receptive language
4.	FTII	Items for testing at 27, 29, 39 and 52 postnatal weeks		Measures visual novelty preference/ recognition memory
5.	WPPSI-III	2 years 6 months–7 years 3 months: (a) 2 years 6 months–3 years 11 months (b) 4 years 0 months–7 years 3 months	45 min for core six subtests	(a) IQ scores + general language composite available (b) Processing speed quotient in addition to above
6.	WISC-IV	7 years 4 months–16 years 11 months	Varies by number of subtests. Approximately 90 min in full	VIQ, PIQ and FSIQ scores + four index scores
7.	WASI	6–89 years	30 min for four subtests	VIQ, PIQ and FSIQ
8.	KABC-II	3–18 years	25–70 min	Scores can be calculated to exclude verbal ability if appropriate. Scales: simultaneous processing, sequential, planning. Learning, and knowledge (if appropriate)
9.	BAS2	Early years: 2 years 6 months–7 years 11 months School age: 6 years–17 years 11 months	Varies by form and number of subtests	General conceptual ability score (GCA) and three embedded cluster scores from six subtests. Further scales available for more specific information
10.	DAS-II	2 years 6 months–17 years 11 months	Core 45–60 months Diagnostics 30 min	Often used to provide basis for intervention in learning problems
11.	CAT3	7 years 6 months–17+ years	45 min/section	Group administration. Three sections: verbal, nonverbal and numerical reasoning. Often used in school settings in UK
12.	Leiter-R	2–21 years	25–40 min	Completely nonverbal. Used to determine if low achievement is due to low IQ or more specific causes
13.	RPM	Colored PM—children Standard PM—general population Advanced PM—top 20% of population	30 min approximately	Group and individual administration. Nonverbal stimuli and simple oral instructions
14.	SON-R	5 years 6 months–17 years	Average of 90 min (45–60 for short form)	Nonverbal. Measures four areas of cognition. Short-form available
15.	WNV	(a) 4–7 years (b) 8–21 years 11 months	30–45 min for four subtests	Native language irrelevant to performance
II—Multi cognitive domains				
16.	NEPSY-II	3–4 years 5–16 years	Core 45 min preschool, 1 h school age. Diagnostics available	Measures six cognitive domains. Sub-tests can be used separately

Table 2 Continued

Level of assessment	Test name	Age range	Time to administer	Notes
17.	CANTAB	Adult and child norms	Tailored to research needs so varies considerably	Computer presentation. Widely used in research. Sensitive to subtle cognitive changes
18.	CDR	Adult test but could be adapted for use with children		Computer presentation. Usually used in a commercial setting—little academic research.
III—Individual cognitive domains				
Memory				
19.	CMS	5–16 years	30 min	Six core sub-tests give index scores, including general memory
20.	TOMAL	5 years–19 years 11 months	45 min	Ten core, four supplementary subtests
21.	WRAML2	5–90 years	45–60 min	Six core sub-tests give three index scores that give general memory quotient
22.	(a) RBMT-II (b) RBMT-C	Use over age of 11 years 5 years–10 years 11 months	Approximately 30 min	Four parallel versions available. Emphasis is on memory in everyday situations not the laboratory
23.	Bus Story Test	3–8 years	15 min	Designed as a language test but can be used to measure narrative memory
24.	CAVLT-2	6 years 6 months–17 years 11 months	25 min	Word list-learning format from which a wide variety of scores is obtained
25.	VADS	5 years 6 months–12 years	10 min	Varied measures of short-term memory
26.	RCFT	6–89 years	45 min, including 30 min delay interval	Measures visuospatial ability and memory. Developmental scoring systems available
Language				
27.	MacArthur-Bates Communicative Development Inventories (a) CDI words and gestures (b) CDI words and sentences (c) CDI-III	8–16 months 16–30 months 30–37 months		Checklist format, parental report well-validated against other data. Many language versions available
28.	CELF-preschool 2 CELF-4	3–6 years 5 years–16 years 11 months	30–45 min 30–60 min	Primary measures of receptive and expressive language plus subtests
29.	BPVS-II	3 years–15 years 8 months	5–8 min	Measures size of the child's receptive lexicon
30.	TROG-2	4-year adult	10–20 min	Measures how well grammatical constructs are understood using a four picture multiple-choice format
31.	Word finding vocabulary test	3–9 years	15 min	50 picture cards for naming
Attention				
32.	TEA-Ch	6–16 years	55–60 min	Five aspects of attention, no overall quotient. Separate norms for boys and girls
33.	(a) CPT-II, Version 5 (b) K-CPT	6-year adult 4–5 years	14 min 7 min	Computer presentation for both
34.	TOVA	4– 80 years	25–30 min	Two versions: visual (TOVA) and Auditory (TOVA-A)

Table 2 Continued

Level of assessment	Test name	Age range	Time to administer	Notes
Executive functions				
35.	BRIEF-PBRIEF	2 years–5 years 11 months 5–18 years	10–15 min	Rating scales. BRIEF has forms for parents and teachers
36.	BADS-C	7–16 years	35–45 min	
37.	CCTT	8–16 years	5–7 min	Three additional forms available for use in research studies
38.	Stroop	5–14 years	10–15 minutes	
Visual functions				
39.	WRAVMA	3–17 years	5–10 min	
40.	Beery VMI	Full form; Short form; Two supplementary tests	15 min; 10 min; 5 min each	
41.	BJLO	Adult test but has some child norms	15 min	Part of the Benton Laboratory of Neuropsychology

Appendix 3: Notes on Websites

The British Psychological Society operates a website that includes details of a wide range of tests and gives access to reviews of them, along with discussion of a range of issues associated with psychological testing. We have also included a list of some major test suppliers' websites; you can search the websites for specific tests by name. Some suppliers require professional eligibility and registration for buying certain tests, so that they may only be available for psychologists, speech therapists, etc., but this information will be available on the website. There are many smaller suppliers so we hope this list is useful, but it cannot be exhaustive and an internet search engine will prove invaluable.

(a) Weblink to British Psychological Society website for psychological testing:

http://www.psychtesting.org.uk/

(b) Weblinks to test suppliers:

Association of Test Publishers
http://www.testpublishers.org/memserv.htm
Cambridge Cognition (CANTAB)
http://www.cantabeclipse.com
Cognitive Drug Research (CDR)
http://www.cognitivedrugresearch.com
Educational and Industrial Testing Service
http://www.edits.net
ERIC Test Publisher Directory
http://buros.unl.edu/buros/jsp/search.jsp
Harcourt Assessment
http://harcourtassessment.com
nferNELSON
http://www.nfer-nelson.co.uk
Pearson Assessments
http://ags.pearsonassessments.com/
Pro-Ed, Inc.
http://www.proedinc.com
Psychological Assessment Resources Inc.
http://www.parinc.com
Taskmaster
http://www.taskmasteronline.co.uk
Western Psychological Services
http://www.wpspublish.com

References

1. Armitage P, Berry G, Matthews JNS (2002) Statistical methods in medical research, 4th edn. Blackwell Sciences, Oxford
2. Baddeley A (1990) Human memory: theory and practice. Lawrence Erlbaum Associates, Hove and London
3. Benton D (2001) Micro-nutrient supplementation and the intelligence of children. Neurosci Behav Rev 25:298–309
4. Benton D, Kallus KW, Schmitt JAJ (2005) How should we measure nutrition-induced improvements in memory? Eur J Nutr 44:485–498
5. Benton D, Ruffin HP, Lassel T, Nabb S, Messaoudi M, Vinoy S, Desor D, Lang V (2003) The delivery rate of dietary carbohydrates affects cognitive performance in both rats and humans. Psychopharm (Berlin) 166:86–90
6. Bornstein MH, Sigman MD (1986) Continuity in mental development from infancy. Child Dev 57:251–274
7. Caspi A et al (2007) Moderation of breastfeeding effects on the IQ by genetic variation in fatty acid metabolism. PNAS Early edition, 0704292104
8. de Graff Peters VB, Hadders-Algra M (2006) Ontogeny of the human central nervous system: what is happening when? Early Hum Dev 82:257–266
9. Deary IJ (2000) Looking down on human intelligence. Oxford University Press, Oxford
10. Diamond A (2002) Normal development of prefrontal cortex from birth to young adulthood: cognitive functions, anatomy, and biochemistry. In: Stuss DT, Knight RT (eds) Principles of frontal lobe function. Oxford University Press, New York, pp 466–503
11. Dobbing J (1981) The later development of the brain and its vulnerability In: Davis JA, Dobbing J (eds) Scientific foundations of paediatrics, 2nd edn. Heinemann, London
12. Dobbing J, Sands J (1973) Quantitative growth and development of human brain. Arch Dis Child 48:757–767
13. Elardo R, Bradley R (1981) Home observation for measurement of the environment: a review of research. Dev Rev 1:113–145
14. Elman JL, Bates EA, Johnson MH, Karmiloff-Smith A, Parisi D, Plunkett K (1998) Rethinking innateness. MIT Press, Cambridge/London
15. Epstein HT (1978) Growth spurts during brain development: implications for educational policy and practice. In: Chard JS, Mirsky AF (eds) Education and the brain (NSSE Yearbook, pt. 2). University of Chicago Press, Chicago, pp 343–371
16. Flynn JR (1984) The mean IQ of Americans: massive gains 1932 to 1978. Psychol Bull 95:29–51
17. Geisinger K (1994) Cross-cultural normative assessment: translation and adaptation issues influencing the normative interpretation of assessment instruments. Psychol Assess 6:304–312
18. Greenwood CE, Craig REA (1987) Dietary influences on brain function: implications during periods of neuronal maturation. Curr Topics Nutr Dis 16:159–216
19. Hebb DO (1966) A textbook of psychology, 2nd edn. WB Saunders Co., Philadelphia/London
20. Herschkowitz N (1988) Brain development in the fetus, neonate and infant. Biol Neonate 54:1–19
21. Horwood LJ, Darlow BA, Mogridge N (2001) Breast milk feeding and cognitive ability at 7–8 years. Arch Dis Child 84:F23–F27
22. Kallus KW, Schmitt JAJ, Benton D (2005) Attention, psychomotor functions and age. Eur J Nutr 44:465–484
23. Kline P (1996) Intelligence: the psychometric view. Routledge, London/New York
24. Kressley RA, Knopf M (2006) A comparison of between- and within-subjects imitation designs. Infant Behav Res 29:566–573
25. Lezak MD, Howieson DB, Loring DW, Hannay HJ, Fischer JS (2004) Neurospsychological assessment, 4th edn. OUP, New York/Oxford
26. Mirsky AF, Anthony BJ, Duncan CC, Ahearn MB, Kellam SG (1991) Analysis of the elements of attention: a neuropsychological approach. Neuropsychol Rev 2:109–145
27. Nettlebeck T (2001) Correlation between inspection time and psychometric abilities: a personal interpretation. Intelligence 29:459–474
28. Nunnally J, Bernstein I (1994) Psychometric theory, 3rd edn. McGraw Hill, New York
29. Paus T, Collins DL, Evans AC, Leonard G, Pike B, Zijdenbos A (2001) Maturation of white matter in the human brain: a review of magnetic resonance studies. Brain Res Bull 54:255–266
30. Perret-Clermont A-N, Carugati F, Oates J (2004) A socio-cognitive perspective on learning and cognitive development. In: Oates J, Grayson A (eds) Cognitive and language development in children. Blackwell, Oxford, pp 303–332
31. Posner MI, Petersen SE (1990) The attention system of the human brain. Annu Rev Neurosci 13:25–42
32. Sameroff AJ, Mackenzie MJ (2003) Research strategies for capturing transactional models of development: the limits of the possible. Dev Psychopathol 15:613–640
33. Sattler JM (1994) Assessment of children: cognitive applications, 4th edn. Jerome M. Sattler, Publisher, Inc., San Diego
34. Schmitt JAJ, Benton D, Kallus KW (2005) General methodological considerations for the assessment of nutritional influences on human cognitive functions. Eur J Nutr 44:459–464
35. Toga AW, Thompson PM, Sowell ER (2006) Mapping brain maturation. Trends Neurosci 29:148–159
36. Westenhoefer J, Bellisle F, Blundell JE, de Vries J, Edwards D, Kallus W, Milon H, Pannemans D, Tuijtelaars S, Tuorila H (2004) PASSCLAIM mental state and performance. Eur J Nutr 43(Suppl 2):II/85–II/117
37. Willatts P, Forsyth JS (2000) The role of long-chain polyunsaturated fatty acids in infant cognitive development. Prostoglandins Leukot Essent Fatty Acids 63:95–100