Assessing Neurocognitive Development in Studies of Nutrition

John Colombo

Department of Psychology and Schiefelbusch Institute for Life Span Studies, University of Kansas, Lawrence, KS, USA

Abstract

As research on clinical nutrition has become more concerned with the effects of macro- and micronutrients on cognitive and brain development, success in evaluating and interpreting those effects is critically dependent on how human cognitive development is conceptualized and measured. The body of research on neurocognitive development from the past 50 years indicates that various cognitive components are relatively independent of one another and develop at different times during infancy and early childhood. For many studies in this area, however, the choice of measures of cognitive development for inclusion in clinical trials has not been guided by a particular theory of cognition or on the hypothesized effect of the nutrient. This practice is potentially disadvantageous for the interpretation of studies in the field; studies may choose neurocognitive assessments which may either obscure the specific effects of a particular nutrient or miss such specific effects altogether because the appropriate domain was not assessed. In developmental studies, this complex scenario is further compounded by the consideration of age-appropriate assessments and domains. This chapter will describe the difficulties in choosing and interpreting cognitive assessments for this field and make recommendations for best practices in addressing this issue.

Introduction

Over the past 2 decades, research on clinical nutrition has become increasingly focused on the effects of macro- and micronutrients on brain development and cognition [1–3]. This is particularly true for studies of infancy and early child-
hood, since it has been hypothesized that nutritional supplementation during
the early period of life may exert long-lasting and meaningful influences on
brain structure and function [4]. Yet, this relatively new focus in the field of
clinical nutrition poses a challenge for the field in the consideration of how neu-
rodevelopment should be measured [5]. For the most part, researchers have had
two broad choices in this realm: (a) standardized tests of global neurodevelop-
ment or (b) nonstandardized assays of specific aspects or *domains* of neurode-
velopment [6]. The former tests have well-defined protocols, can be adminis-
tered without specific technical equipment or expertise, are widely used by (and
well known to) health care practitioners, and are relatively easy to interpret. The
latter typically involve specific cognitive or behavioral tasks that may require
complex protocols, specialized laboratory equipment, or extensive post collec-
tion coding; they are not very well known or used by clinicians, and they may be
difficult to interpret.

Given this comparison, it is not surprising that the field has often made the
decision to use standardized tests (e.g., the Bayley Scales of Infant Develop-
ment, the Griffiths Scales, or the Mullen Early Learning Scales) for evaluating
neurodevelopment in infants and toddlers. At face value, given the desire for
large trials in the context of dwindling resources, such decisions are reasonable.
However, such choices can be disadvantageous if these measures are not sensi-
tive to the specific effects of the nutrient(s) under investigation. Indeed, some
compilations of studies using such measures have yielded conclusions that nu-
trients such as iron [7], zinc [8], or various forms of long-chain polyunsatur-
ated fatty acids (LC-PUFA) [9, 10] convey little or no benefit in the domain
of early neurocognitive development. In the face of compelling theory, mecha-
nisms, and animal models supporting the putative positive effects of these nu-
trients on neurocognitive development, it seems worthwhile to raise the ques-
tion of whether these measures are appropriately sensitive to nutritional ef-
effects.

Indeed, the proper characterization of cognitive development and the use of
appropriate and valid measures of that construct are critical to the evaluation
and interpretation of the effects of nutrition. However, for many studies in this
area, the choice of neurocognitive measures has not been guided by a careful
theory of measurement or by the hypothesized effect of the nutrient [6]. In de-
velopmental studies, the complexity of the choice of cognitive outcomes is com-
pounded by the consideration of which assessments and domains are appropri-
ate at different ages. The objective of this chapter is to provide a modern con-
ceptualization of neurocognitive measurement for nutrition scientists and
practitioners and to make recommendations for best practices for choosing and
interpreting cognitive assessments for this field.
Global and Modular Cognition: Underlying Models

Global neurocognitive assessments are generally based on a unitary model of cognition, in which intelligence may be generally characterized in terms of a single factor [11] from which ability or capacity in all cognitive domains is thought to derive (Fig. 1a). This general factor is often thought to be mediated...
by basic biological parameters of brain function that permeate all cognitive operations and is often attributed to genetic factors [12]. Given this model, it makes some sense for neurocognitive performance to be characterized in terms of a single number; a single composite measure representing cognitive performance is intuitively appealing and highly convenient for analysis and interpretation.

However, the validity of the general model rests on a critical but questionable assumption. Intelligence tests are typically structured such that an overall composite score (e.g., an intelligence quotient or IQ) is computed from a combination of subtests or subscales that measure specific domains of knowledge or ability. A unitary or general model of cognition predicts that these subdomain scales or subtests should be highly correlated with one another. As it turns out, the subscales of IQ tests are only modestly correlated; they share only 10–25% of overall variance [13]. If considered from the other side of the argument, 75–90% of the variability in IQ is attributable to skills that reflect specific functions or domains. Thus, while IQ subscales are, therefore, not completely independent from one another, they are also not highly correlated with one another. Thus, in choosing an overall or composite outcome in a nutrition-related clinical trial, one is gambling that the effect of the nutrient involved will either be represented in the small amount of shared variance among the subtests, or that the effect will not be obscured by aggregating more sensitive subtests or items into the composite score with less sensitive subtests or items.

Historically, the single-factor model can be traced to the first half of the 20th century. More modern conceptualizations of cognitive function derived from information processing theory [14] and neuroscience [15, 16] suggest a more modular model of cognition [17] in which cognitive performance can be conceptualized in terms of the operation of distinct, independent, and specific modules that map onto specific domains [18] such as attention, working memory, long-term memory, language, and executive function (Fig. 1b).

The utility of assessing specific domains has been demonstrated in studies of prenatal teratogens. For example, prenatal exposure to alcohol produces postnatal deficits in visual attention and reaction time [19] but memory is left intact. However, infants prenatally exposed to polychlorinated biphenyls show deficits in memory, but not in attention or speed of processing [20]. Thus, if different nutrients or micronutrients affect different cognitive systems, outcome measures will need to be selected carefully, and global tests may obscure specific effects. Indeed, neuroscience-based research has shown that even moderate levels of granularity may be inadequate; consider recent work showing that long-term memory alone is not a unitary construct and is actually comprised of no less than 6 different functions, each associated with its own underlying neural system.
Similar analyses suggest that attention reflects multiple operations mediated by numerous brain pathways and their interactions.

In summary, choosing a global composite outcome for studying the effects of nutrients in clinical trials is a risky proposition. A more granular and domain-specific approach to the measurement of neurocognitive development seems more desirable, especially if one suspects that the effect of a nutrient might be specific to some neurocognitive functions but not others.

**Developmental Implications**

Aside from the debate over the appropriateness of global neurocognitive tests versus more granular assays, the emphasis on measuring early neurocognitive development presents yet another challenge for the field. Said more simply, once one has decided what to measure, the question arises as to when it might be best to measure it. Once again, at face value, this does not appear to be a difficult decision if one has chosen a global developmental assessment, as standard composites or subscales can be readily derived for different ages.

However, infancy and early childhood is a time of great and rapid change in brain systems and behavioral repertoire, and what is critical to measure at one time during this period is not what might be critical to measure at others. This truth is obvious when one has chosen more modular/granular outcomes. For example, if one is assaying neurodevelopment during the neonatal period, the quality or distribution of sleep or perhaps heart rate variability (a psychophysiological index of CNS integrity) might be excellent candidates, as these variables are undergoing emergent and rapid development. However, just a few weeks later, one might choose a very different matrix of outcomes as sleep patterns are entrained and consolidated by then, and CNS-mediated physiology has settled and stabilized. What is not commonly noted, however, is that this principle is manifest within global developmental tests as well: the set of items administered in (for example) the Bayley Scales of Infant Development at 2 months of age is almost entirely different from the set of items administered at 6 months of age. Thus, whether one uses global developmental tests or more granular neurocognitive tasks, the focus of measurement is shifted so that biobehavioral systems are assayed when individual differences in those systems are most variable and meaningful, i.e., when those systems are emergent or developing at very rapid rates. Thus, at the earliest ages, we might assess systemic indices of basic vital functions or sensory development. In early to mid-infancy, we might assay simple lower-order cognitive functions, such as attention and memory. In later infancy, we would measure higher-order (regulated) abilities that reflect the inte-
gration and coordination of lower-order components [22, 23] that yield the capacity for behavioral inhibition and rule learning and retention that support simple goal-directed behavior. Finally, in early childhood, we would measure higher-order abilities in challenging strategic or adaptive contexts, such as cognitive flexibility or problem solving.

**General and Modular Models of Cognition in a Developmental Context**

An additional issue in the choice of measurement strategies for nutrition clinical trials is revealed by contrasting the general/unitary cognition and modular cognition models from a developmental perspective. A critical issue in psychological assessment is the degree to which a measure accurately reflects individual differences on an ability or skill; among the issues facing developmental scientists is whether that ability or skill is consistent or stable across time. A general cognition model holds that continuity in neurocognitive development over time would be attributable to the general cognition factor; since, in this mode, the general factor drives individual differences in specific measures, then one needs to not be particularly careful or critical regarding the choice of outcome measures across development (Fig. 2a). If all domains are driven by the same underlying factor, then any presumed continuity across time should exist across domains. This general scheme is called *homotypic continuity*.

With a modular model of cognition, however, continuity may be conceptualized more as a developmental cascade than as a direct path across time. Here, different components emerge at different times during development, and these different components become integrated or coordinated to yield higher-order or more sophisticated forms of cognition (Fig. 2b). Under this model, one might expect that some components would contribute variance to some degree to continuity in developmental outcomes but through indirect paths. Evidence for this *heterotypic continuity* has been borne out in analyses seeking to determine how, for example, early individual differences in attention contribute to later cognitive and language outcomes in childhood [24, 25].

This perspective reinforces the point that different measures should be taken in evaluating the neurocognitive effects of nutrients at different times during infancy and early childhood. However, it also raises the point that changes to lower-order neurocognitive components (e.g., attention and memory) early in development can produce important changes in more complex neurocognitive components (e.g., executive function and language) later in childhood or the school-age years. Given the dynamic nature of change across the early part of the life span, it may be important to track the development (i.e., change) of these neurocognitive components in clinical trials. We turn our attention to this point briefly in the next section of the chapter.
Fig. 2. This figure shows the models of neurocognition outlined in Figure 1 but in a developmental context.  

**a** The general model of cognition or intelligence; since all neurocognitive components are driven by a single underlying factor, continuity across development is mediated by that general factor, and all subcomponents emerge as manifestations of that factor, showing similar degrees of interrelatedness.  

**b** The modular model where, over time (right to left), lower-order subcomponents (e.g., attention and memory) become integrated or coordinated to influence or yield higher-order components.
Evaluating Developmental Course

As noted above, a common question asked in the design of clinical trials is when to measure certain outcomes. A less common question is how often those measures should be taken. Early on, nutrition studies were content to assess the effects of nutrient status and supplementation on neurocognition with a single outcome. Indeed, the conclusions of meta-analyses cited previously were based largely on global tests measured at a single time point. However, this “snapshot” approach to measurement misses the opportunity to examine the developmental course of neurocognitive functions which, in the context of the heterotypic cascade represented in Figure 2, may provide added sensitivity to the design of clinical trials.

We have made the theoretical case elsewhere [5, 26] for measuring the developmental course of neurocognitive outcomes in clinical trials, but we can now provide multiple specific examples to bolster that case. Figure 3a is derived from a randomized clinical trial of zinc supplementation [27] conducted on an iron- and zinc-deficient sample in Peru; we provided iron to all infants (thus eliminating anemia in the sample) but provided zinc to half the sample. We hypothesized that zinc would restore normative neurocognitive development, a prediction borne out by our measurement of visual attention during the 1st year. Typically, during the 1st year, infants’ duration of looking to a visual stimulus declines, because they become faster at processing and encoding information. Indeed, providing zinc resulted in the typical normative decline in look duration, while the placebo group showed no change across the 1st year. Note that the two groups shown in Figure 3a vary only on the last data point (at 12 months); if we had taken a “snapshot” evaluation of this outcome at either of the two previous ages, we would have concluded that the supplementation had no effect.

Another example is from a study of the long-term effects of 12 months of feeding infant formula supplemented with LC-PUFA docosahexaenoic and arachidonic acid on neurocognitive outcomes through 5 years of age [28]. Here, we took multiple assessments of the Dimensional Change Card Sort, a measure of emergent executive function; the task was administered at 3, 3.5, 4, and 5 years of age. The data show (Fig. 3b) that the supplemented and unsupplemented groups did not vary at 36 or 42 months. However, infants fed the supplemented formula showed significant improvement in the task beginning at 4 years of age, while infants fed a formula with no LC-PUFA did not improve on the task until 5 years of age; thus, the nutritional supplementation accelerated the development of early executive function by 1 full year. Again, the key to demonstrating this striking effect was the measurement of the developmental course on this outcome at a sensitive time of change.
In summary, the discussion presented here leads to a number of recommendations for the design and implementation of clinical trials in nutrition where the outcome is neurodevelopment, and for health care practitioners who may be consumers of the extant literature on the effects of nutrition on neurodevelopment.

**Design and Implementation of Clinical Trials**

Fundamentally, the choice of global developmental tests for the assessment of neurodevelopment is presented as an extremely risky one. Global developmental tests present aggregate, composite scores that presumably represent an overall assay of neurocognitive performance. While it is certainly possible that certain micro- or macronutrients may produce effects large enough to be detected with these outcomes, if there is any reason to believe that the nutrient effect may

---

**Fig. 3.** This figure shows the value of tracking the developmental course of neurocognitive components in clinical trials. 

**a** The development of look duration for infants supplemented with iron (dashed line) or zinc plus iron (solid line); the normal course of development is a decline with age, and the decline is only seen when both nutrients are provided. **b** The development of a simple measure of executive function (rule learning and flexibility) from a study in which infants received a formula with no LC-PUFA (placebo) or formula supplemented with various levels of LC-PUFA; note that the placebo group shows no improvement until 5 years of age, while improvement is seen in supplemented infants 1 full year earlier.
be subtle, or (more importantly) may be specific to certain neurocognitive systems (e.g., attention or memory) or subsystems (e.g., declarative memory, inhibition, or attentional regulation), then there is a good probability that global developmental tests will miss them.

If more granular tests are chosen for inclusion on clinical trials, then the choice of outcome should be informed by the mechanisms through which the nutrient under question is presumed to affect neurocognitive outcomes, or the specific systems through which those effects are expected to be manifest. The design of the trial should involve at least a short-term longitudinal measurement strategy, in which the choice and timing of outcome measurement should be either emergent or developing rapidly. If possible, the outcome measures should be collected at least at 2–3 points during that time of rapid change and maturation. A template for constructing assessment schedules in clinical trials (Fig. 4) should follow the general principles of including increasingly sophisticated (but developmentally appropriate) tasks across age, which are administered at least twice (and preferably 3 times) for optimizing sensitivity.
**Reading the Nutrition and Neurodevelopmental Literature**

A final recommendation from this chapter is directed to consumers of the nutrition literature, including journal editors, scholars in the area, and policy makers for early nutrition. A century of research on cognitive and neurocognitive development has taught us a number of lessons, and among those are that neurocognitive development is complex and dynamic, and small changes to lower-order cognitive components early in life can compound over time to affect increasingly complex behaviors later on. Thus, the sensitivity of assessments in evaluating the effects of nutrients on neurocognition is a paramount consideration in interpreting the outcome of clinical trials. The use of global developmental tests tends to aggregate performance from a number of neurocognitive domains or components and, in that aggregation, subtle but potentially important effects might be obscured. Thus, null findings derived from trials employing only such global tests or using limited windows of assessment (e.g., single-visit “snapshot” studies) should be viewed with caution, even when the sample sizes might be large. The effects of nutrients are best assessed with rich longitudinal schedules that allow for the tracking of developmental functions.

**Disclosure Statement**

Preparation of this chapter was supported in part by US National Institute of Health grants U54HD090216, R01HD086001, and R01HD047315. The author is an occasional consultant to Nestle Research Center.

**References**

5 Colombo J: Recent advances in infant cognition: implications for long-chain polyunsaturated fatty acid supplementation studies. Lipids 2001;36:919–926.
8 Black MM: Micronutrient deficiencies and cognitive functioning. J Nutr 2003;133: 3927S–3931S.


13 Detterman DK, Daniel MH: Correlations of mental tests with each other and with cognitive variables are highest for low IQ groups. Intelligence 1989;13:349–359.


