An overview of evidence for a causal relation between iron deficiency during development and deficits in cognitive or behavioral function

Joyce C McCann and Bruce N Ames

ABSTRACT
This review, intended for a broad scientific readership, summarizes evidence relevant to whether a causal relation exists between dietary iron deficiency with (ID+A) or without (ID-A) anemia during development and deficits in subsequent cognitive or behavioral performance. An overview of expert opinion and major evidence in humans and animals is provided. Cognitive and behavioral effects observed in humans with ID-A and in animals with ID±A are provided in tables. The degree to which 5 conditions of causality are satisfied and whether deleterious effects of ID-A might be expected to occur are discussed. On the basis of the existing literature, our major conclusions are as follows. Although most of the 5 conditions of causality (association, plausible biological mechanisms, dose response, ability to manipulate the effect, and specificity of cause and effect) are partially satisfied in humans, animals, or both, a causal connection has not been clearly established. In animals, deficits in motor activity are consistently associated with severe ID+A, but adverse effects on performance in tests that target cognitive function have not been clearly shown. Resistance to iron treatment was observed in most trials of children <2 y of age with ID+A, but not in older children. Similar observations were made in rodents when ID+A occurred before rather than after weaning. In children >2 y of age and in adolescents with ID-A, evidence suggests cognitive or behavioral deficits; however, the surprisingly small number of studies conducted in either humans or animals prevents a thorough assessment. Am J Clin Nutr 2007;85:931–45.

KEY WORDS   Iron, anemia, iron deficiency, cognition, behavior, learning, memory, gestation, pregnancy, brain, neurology, infants, childhood, rodent studies

INTRODUCTION
A large body of research suggests that an inadequate dietary supply of any of a number of essential micronutrients can adversely affect brain function (1–6). Some studies also suggest positive effects of multivitamin and mineral supplementation on cognitive function (7, 8). The brain is at its most vulnerable during critical periods of development, including the last trimester of fetal life and the first 2 y of childhood—a period of rapid brain growth termed the “brain growth spurt” (9). This review is part of a series intended to provide critical summaries of the available experimental evidence pertinent to whether causal linkages exist between individual micronutrient deficiencies during this critical period and subsequent brain function. We recently reviewed evidence of a causal relation between cognitive dysfunction and n–3 fatty acid deficiency (5) or choline availability (6) during development.

A causal relation between micronutrient deficiencies and suboptimal brain function would have major public health implications. Large segments of the world (including the United States) population, particularly the poor, are known to be undernourished in a number of micronutrients (10–12). A major effort to address micronutrient undernutrition as an adjunct to the various programs underway to improve dietary habits, particularly of the poor, will be well justified. One of us has discussed such an approach as a relatively inexpensive and efficacious adjunct to current public health programs (13, 14).

Dietary iron deficiency sufficient to cause anemia (ID+A) is widespread in underdeveloped countries (15, 16). Even in countries such as the United States, the prevalence of ID+A is as high as 29% among low-income pregnant women (17, 18); iron intake is below the Estimated Average Requirement (EAR) for ≈16% of menstruating women (19). The National Health and Nutrition Examination Survey (NHANES) of 1999–2000 reported that, among all age groups examined, the estimated prevalence of ID±A was greatest in adolescent girls (9–16%) and in young children during the period of the brain growth spurt (7%) (20, 21). Worldwide, the economic effect of ID+A has been estimated to be in the billions of dollars (22).

Although this review focuses on dietary iron deficiency, there are other causes of iron deficiency in children. Infants subjected to conditions of pregnancy resulting in intrauterine growth restriction (IUGR), infants who are small-for-gestational age (SGA), infants of diabetic mothers (IDMs), or infants born of preeclamptic mothers can also be iron deficient (23–27). The reader is referred to a large body of literature describing cognitive or behavioral deficits in such children; several reviews and examples are cited (28–34). Other factors associated with iron deficiency...
deficiency in infants are early umbilical cord clamping, prematurity, and fetal blood loss (35).

We present an overview of evidence relevant to establishing whether a causal link exists between dietary iron deficiency during development and subsequent cognitive or behavioral function. The degree to which this evidence satisfies 5 causal criteria, slightly modified from the original formulation (36), is examined: 1) a consistent association, 2) a dose-response relation, 3) ability to manipulate the effect, 4) specificity of cause and effect, and 5) a plausible biological rationale. To address these criteria, it is essential to consider evidence from both human and animal systems. Animal studies provide important corroborative evidence and other information that is difficult or impossible to obtain in human studies. For example, some types of experiments cannot be done in humans, such as mechanistic studies that correlate changes in biochemical indicators of brain function to performance deficits on cognitive or behavioral tests. Also, much greater flexibility of design is possible in animal experiments, which avoids some serious confounders in human studies. Thus, this review addresses an issue of importance to pediatrics by taking into account the full array of relevant scientific evidence from both human and animal systems.

An in-depth methodologic review of this large body of evidence is beyond the scope of this review; systematic critical reviews were relied on to the extent possible. There are several such reviews of human studies (1, 37, 40), but, to our knowledge, a systematic critical review of cognitive and behavioral studies in iron-deficient animals has not been conducted. We searched the literature using a combination of techniques, including key word and author searches, using the National Library of Medicine’s PUBMED and Science Citation Index Cited References databases; we also surveyed citations included in recent research and review articles. Abstracts were not included.

**Background**

Iron is required for many essential bodily functions, including oxygen transport, ATP production, DNA synthesis, mitochondrial function, and protection of cells from oxidative damage, as discussed in many reviews (41–44). The average concentration of iron in the brain is far higher than that of all other metals, except zinc (45). As widely reviewed (1, 46–49), iron is required by enzymes involved in specific brain functions, including myelination (50–52) and synthesis of the neurotransmitters serotonin (tryptophan hydroxylase) (53) and dopamine (tyrosine hydroxylase), a precursor to epinephrine and norepinephrine (54).

**Accretion of iron by the brain**

Accretion of iron by the human fetus begins early in pregnancy (55), increases dramatically in the third trimester (56, 57), and continues after birth up to 30–50 years of age (58, 59). Unless iron-deficient anemic (ID-A) (15, 36), infants (tryptophan hydroxylase) (53) and dopamine (tyrosine hydroxylase), a precursor to epinephrine and norepinephrine (54).

**Performance tests**

Standardized tests that screen broadly for cognitive and related functions, particularly the Bayley Scales of Infant Development (72), have been most commonly used in children <2 years of age, although some tests that target cognitive function more specifically, such as the Fagan Test of Infant Intelligence (73), have also been used. In older children, a broader range of tests have been used, including the Stanford-Binet Intelligence Scale (74), the Wechsler Intelligence Scale for Children (WISC) (75), several attentional tests (eg, 76), and school achievement tests. A few human experiments also used electrophysiologic measures (77–79). For a detailed listing of tests used in many human studies, the reader is referred to the excellent review of Grantham-McGregor and Ani (38).

In the 2 nonhuman primate studies identified, a wide range of neurobehavioral assessments were used, some of which are similar to human tests, such as an adaptation (80) of the Fagan Test of Infant Intelligence (73). In rodents, motor or exploratory activity was most frequently measured, using activity monitors, the hole board, or home orientation and open-field tests. However, some methods that target cognitive functions, such as learning and memory, more directly, such as the Morris water maze (81) and passive or active avoidance tests (82), were also used. For further examples and discussion of rodent performance tests, see the article by Metz et al (83).

**Definition of iron deficiency with and without anemia**

Sensitivity to iron-restrictive conditions depends on the severity of dietary restriction but also on the stage of development during which iron deficiency occurs and the duration of dietary restriction (84–87). Dietary iron deficiency results in biochemical changes in the blood and reduced concentrations of iron in tissues. ID-A is generally considered to correspond to a degree of dietary iron deficiency sufficient to deplete ferritin stores and to decrease iron concentrations in some tissues, but not sufficient to reduce serum hemoglobin to the point of anemia. Individuals with depleted iron stores and serum hemoglobin concentrations below the 98th percentile of a normally distributed population are generally considered to be iron deficient anemic (ID+A) (15, 44). Hemoglobin cutoff values vary somewhat among human studies, as noted by reviewers (38).

**HUMAN STUDIES**

Many studies in humans have examined possible linkages between iron deficiency (primarily ID+A) and concurrent or subsequent cognitive or behavioral outcomes; most of these studies have been thoroughly reviewed (1, 37–40, 88–98). Although concurrent effects may reflect neurochemical changes resulting from iron deficiency at the time of testing, demonstration that a permanent developmental change has occurred requires evidence of effects in formerly iron-deficient children (40).

Reviewers emphasize the importance of making causal inferences only from studies with designs that minimize potential confounding (1, 38–40, 84, 99). For example, the anemia that accompanies ID+A can be a potentially serious confounder in case-control studies that examine concurrent effects of ID+A. Complex relations between cognitive or behavioral outcomes and socioeconomic factors are also very difficult to completely control for. As the
Collectively, expert reviewers discussed >40 studies of various experimental designs that examined the performance of children or adolescents with ID+A in cognitive or behavioral tests (1, 37–40, 88–97). >60% of the tests were conducted in children with ID+A who were <2 y of age. Cognitive or behavioral tests administered at the beginning of iron-treatment trials provide an indication of the performance of children while ID+A (100–103). For such children, reviewers agreed that there was a consistent association between ID+A and poor performance relative to control subjects (1, 37–40, 90–97). Two comprehensive critical reviews of studies involving children identified as ID+A at ages ≥2 y (38, 89) also concluded that, in most cases, performance was poorer, at least on some tests (104–106). As widely discussed (1, 84, 107–110), it is difficult to link iron deficiency per se to poor test performance. Interpretation is limited, not only by potential confounding due to unaccounted socioeconomic factors, but by the potential confounding factor of anemia that accompanies the iron deficiency of ID+A. In rats, anemia per se does not result in biochemical brain effects associated with ID+A (111). However, it cannot be excluded that the poor performance of ID+A children (or animals) on cognitive or behavioral tests is confounded by generalized effects of anemia on physical energy levels (49, 112–116) as opposed to specific effects of iron deficiency in the brain.

Cognitive or behavioral outcomes in formerly anemic children

Effect of iron treatment on cognition or behavior in children with ID+A

A critical question motivating a great deal of research in the field is whether and how long the effects of ID+A on cognitive or behavioral function persist after children are no longer iron deficient (1, 38, 40, 108). The 2 extensive reviews cited above (1, 38) critically assessed >20 iron-treatment trials involving children with ID+A, most of which had a DBRCT design.

These and other reviewers (1, 37, 38, 117) concluded that, in general, poorer test performance of children with ID+A tended to improve with iron treatment in children ≥2 y of age (104, 118, 119) but was more resistant to improvement in children <2 y of age (101–103). As investigators discussed, this observation is compatible with damage from iron deficiency during brain development being irreversible. Reviewers also pointed out uncertainties associated with this conclusion, such as the limited statistical power of studies that did not observe a treatment effect and an insufficient number of longer-term randomized treatment trials. Only 2 longer-term (2–4 mo) trials reviewed were DBRCTs (120, 121), and 1 of these (120) observed that performance deficits improved with treatment.

The finding that the cognitive performance of children with ID+A at an early age is resistant to improvement with iron treatment is supported by several long-term follow-up studies (38). We point particularly to the widely cited longitudinal observational study of Lozoff et al (40, 122, 123), who followed a group of Costa Rican children for >10 y. In this study, children who formerly had ID+A (initially treated for 3 mo with iron) performed less well than did control subjects when tested at 5 and 11–14 y of age. Extensive test batteries included measures of intelligence quotient, verbal and quantitative learning, memory, and attention. As discussed by the lead investigator (40, 100, 124) and other reviewers (37, 38, 99), although this study provides important evidence, it cannot be considered definitive because it was not a DBRCT and did not have a placebo control group.

Auditory evoked potentials

An outcome measure that avoids the potentially confounding factors of anemia and socioeconomic conditions uses auditory evoked potentials, which are noninvasive electrophysiologic measures of how long it takes the acoustic nerve to transmit sound from the ear to the brain. This time is inversely related to the degree of myelination; the technique is commonly used to detect hypomyelination associated with various diseases (77). Four laboratories used auditory evoked potentials to compare the rate of myelination in control children and children with ID+A (125–130). The results are briefly summarized below.

The most commonly cited studies is that of Lozoff et al (125, 126). In this experiment, conduction times were measured over a 4-y period in Chilean children with ID+A and control subjects. Children were identified as having ID+A or as control subjects at 5–6 mo of age, and all children received iron supplements for 1.5 y. Although results at later time points suggested that the formerly ID+A group was gradually catching up to control subjects, conduction times remained slower at all time points examined. Shankar et al (129) observed a significant correlation between the severity of anemia and conduction time in a group of children with ID+A ranging in age from 3 to 11 y. Several other reports that measured auditory evoked potentials were either inconclusive (130), positive but unavailable for analysis (128), or negative but of uncertain significance because only group means were reported and ages varied within groups across a wide range (7–24 mo) (127).

Preventive trials

Preventive trials aim to determine whether the development of ID+A and developmental deficits can be prevented in hematologically normal children if they are supplemented with iron. The difficulty in obtaining sufficient power in such trials to detect significant effects has been discussed by reviewers (38, 39), who concluded that the few trials conducted provided only limited evidence of benefit.

Two recent trials are noted (131, 132). Friel et al (132) conducted a small (n = 77) DBRCT of a group of breastfed children supplemented with iron or placebo from ages 1 to 6 mo. At 12–18 mo of age, higher visual acuity scores and improved performance on the Bayley Psychomotor Development Index (but not on the Mental Development Index) were observed (132), which led investigators to conclude that supplementation might have beneficial developmental effects. As noted by others (133), the study requires replication with larger group sizes before any definitive conclusions can be made. A second much larger but nonrandomized trial (n > 1000), which examined children at 1 y of age after 6 mo of iron supplementation, did not observe an effect on Psychomotor...
Development Index scores (131). However, longer looking times ($P < 0.01$) and other evidence of developmental behavioral deficits in the nonsupplemented group were observed (131).

Iron deficiency without anemia

Given the significant public health consequences if there are effects of ID-A on cognition or behavior (109, 134), it is surprising that so few studies have focused on this topic. Some 17 investigations in children or adolescents, most of which were conducted >15 y ago (102, 103, 105, 106, 118, 120, 134–144), are listed in Table 1.

As shown, most of the studies were iron-treatment trials that measured cognitive or behavioral performance before and after periods of iron treatment ranging from 1 wk to 6 mo. As discussed by reviewers (38, 94, 145), if there is an effect of ID-A on cognitive or behavioral performance in children <2 y of age, the effect appears to be small. As shown in Table 1, 4 of the 8 treatment trials listed involving children <2 y of age observed deficits in ID-A groups compared with control groups (135–137, 144). However, 3 of these 4 trials (136, 137, 144) were not randomized, and all 4 were quite small. Three of these trials (135–137) were critiqued by some reviewers (38) on other design or quality-control issues. As discussed (93), the 4 trials that observed no differences between ID-A groups and control subjects (102, 103, 120, 138) were also somewhat limited in statistical power, leaving open the possibility that weak effects may occur. As shown in Table 1, a recent report involving newborns (134) is consistent with this possibility. Also suggesting a weak but significant effect of ID-A are all but 1 (105) of the 8 studies listed in Table 1 involving children >2 y of age or adolescents (105, 106, 118, 139–143).

ANIMAL STUDIES

The great strength of animal studies is that they afford the opportunity for more flexibility in design and ability to control experimental variables than can be achieved in human studies. However, important differences exist between animals and humans that must be taken into account in evaluating the implications of results for humans (146). For example, in rats and mice, offspring are born at an earlier stage of development than are humans, with the period of the brain growth spurt generally considered to begin at birth and extend up to weaning (usually postnatal day 21). In humans, the brain growth spurt begins in the last trimester of pregnancy and extends through the first 2 y of life. Another important difference between rodents and humans is that rats have a greater requirement for dietary iron during development than do humans because the rate of development is relatively much more rapid (147). Also, in the great majority of the studies discussed below, the dietary supply of iron was much more severely restricted than in human studies. Thus, species and dosimetry differences must be considered when results in animals are extrapolated to humans.

As for human studies, potential confounders are essential to consider in the design and interpretation of these experiments. In severe iron deficiency experiments, anemia (148, 149) and growth restriction (146, 150, 151) are important potential confounders. These are discussed further in conjunction with specific results below.

Effects on cognitive, behavioral, or motor activity in rodents and monkeys

On the basis of citations in expert reviews (1, 47, 49, 50, 97, 152–167) and on an independent literature search, 20 rodent (48, 85, 168–185) and 2 monkey (186–188) studies that examined cognitive, behavioral, or motor activity in iron-restricted animals were identified (Table 2). As shown, all of these studies reported statistically significant differences between iron-restricted and control groups on at least some tests.

For several variables, the results were surveyed to determine whether differential responses (in degree or type) could be discerned with respect to 1) severe or less-severe iron restriction, 2) dose-response effects, 3) dietary restriction before or after weaning, 4) type of performance test administered, and 5) influence of a period of iron repletion on the test outcome.

Severe compared with less-severe iron restriction

As shown in Table 2, all but 4 of the rodent studies (177–180) used restrictive diets that were ≥90% iron-depleted for at least part of the experimental period (48, 85, 168–176, 179, 181–185). Almost all of these latter experiments included the results of performance tests conducted while the animals were anemic (Table 2, column 3). Possible confounding due to secondary effects of anemia cannot be excluded in these experiments, particularly when animals were severely anemic (Table 2). As also shown in Table 2, animals with ID-A weighed significantly less than did controls (usually 15–20% less). It is known that early malnutrition in rodents can alter later performance on affective and some cognitive tasks (189). It is not clear whether undernutrition in the experiments shown in Table 2, which was relatively modest in most cases, could account for the observed results (146). To minimize potential confounding due to undernutrition, a few investigators used weight controls (169–171, 183) or included body weight as a covariate in statistical analyses (177, 178).

As shown in Table 2, regardless of the degree to which animals were iron-restricted, regardless of whether they were ID-A or ID-A at the time of testing and regardless of the degree of undernutrition or whether it was controlled for, changes in activity or behavior compared with that in controls were reported on at least some tests in all studies.

In our opinion, none of the rodent studies cited above that used somewhat less restrictive dietary conditions (177–180) provide evidence that motor, cognitive, or behavioral deficits result from ID-A (Table 2). Although offspring had normal hematocrit readings at the time of testing by Kwik-Uribe et al (177, 178), they were more severely iron-restricted during gestation, as evidenced by the 50% decrease in hemoglobin on postnatal day 21 and by significant reductions in both body and organ weights, including the brain. As shown in Table 2, the results are mixed and difficult to interpret in the activity study of Hunt et al (179), and hemoglobin decreased ≈40% at the time of testing in the experiment by Massaro and Widmayer (180).

As pointed out by the investigators, in 1 of the 2 monkey studies (186, 187), half of the 8 monkeys in the iron-restricted group were anemic at the only time point that showed a statistically significant ($P < 0.05$) decrease in spontaneous activity. Also, as shown in Table 2, the group restricted in both zinc and iron scored better on one outcome measure on an attentional task and worse on another, which makes interpretation of the results
<table>
<thead>
<tr>
<th>Reference</th>
<th>Study type/treatment duration</th>
<th>ID-A group size/age</th>
<th>Outcome measures</th>
<th>Results reported by investigator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Akman et al, 2004 (144)</td>
<td>Treatment trial6/12/15</td>
<td>40/6–30 mo</td>
<td>DDST, Bayley</td>
<td>Lower test scores than control subjects before or without treatment (DDST-II, MDI) but not after treatment.</td>
</tr>
<tr>
<td>Deinard et al, 1986 (135)</td>
<td>Treatment trial6 mo</td>
<td>45/7.5–5 y</td>
<td>MDI, SB-IQ, Likert, neuropsych</td>
<td>Increase in MDE/SB-IQ scores in control subjects tended to be greater than in the ID-A group, with or without treatment.</td>
</tr>
<tr>
<td>Deinard et al, 1981 (138)</td>
<td>Cross-sectional study/NA</td>
<td>34 or 2/11–13 mo</td>
<td>Bayley, UHSPD, habituation</td>
<td>No significant differences between ID-A and control groups.</td>
</tr>
<tr>
<td>Idradinata and Pollitt, 1993 (120)</td>
<td>Treatment trial4 mo</td>
<td>29/3/12–18 mo</td>
<td>Bayley</td>
<td>No significant differences between ID-A and control groups at baseline.</td>
</tr>
<tr>
<td>Luzzati et al, 1987 (102)</td>
<td>Treatment trial/ wk and 3 mo</td>
<td>38, 21, 45/12–23 mo</td>
<td>Bayley</td>
<td>No significant differences between ID-A and control groups.</td>
</tr>
<tr>
<td>Oski et al, 1983 (137)</td>
<td>Treatment trial2/7 d</td>
<td>28/9–12 mo</td>
<td>Bayley</td>
<td>Greater increase in MDE scores in more severely iron-deficient groups.</td>
</tr>
<tr>
<td>Wachs et al, 2005 (134)</td>
<td>Correlational/NA</td>
<td>24/newborns</td>
<td>Temperament</td>
<td>Several measures of temperament were correlated with hemoglobin or serum iron across a range of severity of iron deficiency that included an unspecified number of ID-A infants.</td>
</tr>
<tr>
<td>Walter et al, 1989 (103)</td>
<td>Treatment trial/10 d and 3 mo</td>
<td>127/1 y</td>
<td>Bayley</td>
<td>No significant differences between ID-A and control groups.</td>
</tr>
<tr>
<td>Walter et al, 1983 (136)</td>
<td>Treatment trial2/15 d</td>
<td>12 or 15/5/15 mo</td>
<td>Bayley</td>
<td>ID-A group scores not significantly different from control subjects before or after treatment; improved MDE scores in an ID-A subgroup.</td>
</tr>
</tbody>
</table>

**Older children or adolescents**

<table>
<thead>
<tr>
<th>Reference</th>
<th>Study type/treatment duration</th>
<th>ID-A group size/age</th>
<th>Outcome measures</th>
<th>Results reported by investigator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bruner et al, 1996 (106)</td>
<td>Treatment trial8 wk</td>
<td>81/7/13–18 y</td>
<td>BTA, SDMT, VSA, HVLT</td>
<td>Greater improvement in HVLT scores after iron treatment in a group of ID-A adolescent girls than in placebo-treated ID-A control subjects.</td>
</tr>
<tr>
<td>Groner et al, 1986 (118)</td>
<td>Treatment trial/1 mo/28</td>
<td>38/14–24 y</td>
<td>WAIS or WISC, CT, RAVLT</td>
<td>Increases in the scores on the arithmetic and digit symbol tests were greater for the ID-A group than for the control subjects.</td>
</tr>
<tr>
<td>Otero et al, 2004 (142)</td>
<td>Cross-sectional/R</td>
<td>28/7–10 y</td>
<td>ERP (oddball)</td>
<td>Altered ERP pattern compared with control subjects; fewer correct answers to oddball paradigm before and after treatment.</td>
</tr>
<tr>
<td>Pollitt et al, 1983 (143)</td>
<td>Treatment trial3 mo</td>
<td>15/2–3/6 y</td>
<td>SB-IQ, Disc Lrn, Odd Lrn</td>
<td>Greater number of trials to reach criterion in Disc Lrn for iron-deficient group before treatment.</td>
</tr>
<tr>
<td>Pollitt et al, 1989 (140)</td>
<td>Treatment trial/16 wk</td>
<td>47/6/1–11 y</td>
<td>Raven C, Educ</td>
<td>Lower scores on language test than iron-replete controls.</td>
</tr>
<tr>
<td>Soewondo et al, 1989 (105)</td>
<td>Treatment trial8 wk</td>
<td>57/2/3/4/5 y (avg)</td>
<td>PPVT, Disc Lrn, Odd Lrn</td>
<td>No differences between ID-A and iron-replete control groups.</td>
</tr>
</tbody>
</table>

1 Bayley, Bayley Scales of Infant Development [includes both the Psychomotor Development Index (PDI) and the Mental Development Index (MDI)]; BTA, Brief Test of Attention; CT, Consonant Trigrams, a test for attention deficits; DDST, Denver Developmental Screening Test; DEL, Dynamic Evaluation of Learning, a computerized learning test; Disc Lrn, Two-choice discrimination learning; Educ, educational achievement tests; EEG, electroencephalogram; ERP, event-related potential, an electrophysiologic measure; HVLT, Hopkins Verbal Learning Test; Likert, Likert Behavioral Rating Scale; Neuropsych, overall neuropsychologic assessment; NHANES, National Health and Nutrition Examination Survey; Odd Lrn, Oddity Learning; PPVT, Peabody Picture Vocabulary Test; Raven C, Raven Colored Progressive Matrices; RCT, randomized controlled trial; RAVLT, Rey Auditory Verbal Learning Test; a test of attention deficits; SDMT, Symbol Digit Modalities Test; SB-IQ, Stanford Binet Intelligence Quotient Scale; UHSPD, 12 girls and Hunt Ordinal Scales (I, II, and V) of Psychological Development; VSA, Visual Search and Attention; WISC-R, Wechsler Intelligence Scale for Children, revised; WAIS, 4 subsets (digit span, digit symbol, arithmetic, and vocabulary) of the Wechsler Adult Intelligence Scale; WRAT-R, Wide-Range Achievement Test, revised.

2 Treatment trials are double-blind, placebo-controlled, randomized trials unless otherwise stated.

3 Studies that also included ID-A groups are indicated in the table.

4 Age at beginning of the treatment period.

5 All results shown are for comparisons between ID-A groups and non-iron-deficient control subjects.

6 Not placebo-controlled.

7 ID-A groups in a larger study.

8 Study critiqued by some reviewers (38).

9 Groups divided according to serum ferritin concentrations: ≤9 ng/mL (n = 34) and 10–19 ng/mL (n = 21); no groups were anemic (hematocrit <34%). This method of allocation was critiqued (94).

10 ID-A children received a placebo during the treatment trial. Reviewers noted the small size of the ID-A group (38).

11 Increasing severity of iron deficiency without frank anemia in the 3 groups. In the most severely affected of the 3 groups (n = 45), hemoglobin values were 10.6–11.9 g/dL.

12 No randomization.

13 Consisted of 3 groups with increasing severity of iron deficiency without frank anemia.

14 183 newborns with iron-status measures ranging from normal to ID-A; the number of ID-A infants was not specified.

15 The P value was based on a one-tailed t test.

16 73 Girls completed the trial; 36 received a placebo.

17 Hematocrit values were >30%; all subjects were in the first 16 wk of pregnancy.

18 Description of results in test and table do not match. Results based on one-tailed t tests.

19 WISC-R tests were digit span and block design subtests; WRAT-R tests were math and reading.

20 Subjects were also supplemented with iron and compared with an unsupplemented control group.

21 Children completed a period of iron treatment and were retested.

22 Children were identified as mildly iron deficient on the basis of a post hoc procedure based on response to iron therapy.

23 No treatment effects observed; results based on averages of pre- and posttreatment scores.

24 Subjects were treated with iron; 33 received a placebo.
### TABLE 2
Cognitive, behavioral, and motor performance tests in animals that are iron deficient anemic (ID+A) or nonanemic (ID-A)

<table>
<thead>
<tr>
<th>References</th>
<th>Diet (% of control)/duration of iron deficiency</th>
<th>Sex/strain</th>
<th>At time of testing/age/iron status (% of control)</th>
<th>Body weight reduction (% of control)</th>
<th>Experimental results for iron-deficient groups compared with controls</th>
<th>Other brain-related experimental measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beard et al, 2002 (48)</td>
<td>3 mg/kg (8.5%)/PND 21–49</td>
<td>d/♀/SD</td>
<td>PND 49/55% (♂), 41% (♀)</td>
<td>60% (♂), 85% (♀)</td>
<td>Light-dark box: more rapid movement into the dark*</td>
<td>Brain iron, dopamine receptor and transporter density</td>
</tr>
<tr>
<td>Golub et al, 2006 (188)</td>
<td>4 mg/kg (10%)/ED 5–PND 7</td>
<td>to PND 20/ED 5–PND 20</td>
<td>PND 6, 9, 12, 15, 18, 21, 24/54% (PND 10, 25)</td>
<td>PND 10, 85%; PND 25, 77% (♀)</td>
<td>Developmental battery: delayed auditory startle, surface righting, negative geotaxis*, bar holding*, bilateral forelimb placing* (PNDs 6, 9, 12, 15, 18, 21); vibrissa-stimulated forelimb placing* (PNDs 12, 15, 18, 21, 24)</td>
<td>Brain iron, monoamines, dopamine D2 receptor, dopamine and serotonin transporters</td>
</tr>
<tr>
<td>Ben-Shachar et al, 1986 (183)</td>
<td>NR/PND 21–49</td>
<td>NR/SD</td>
<td>PNDs 28, 35, 42, 49, 63, 70, 77/60-80%</td>
<td>Food intake of controls restricted to keep weights similar</td>
<td>Activity + apomorphine: fewer movements: PND 35*, 49**, 63% (PND 21-49)</td>
<td>Brain iron, dopamine receptor density</td>
</tr>
<tr>
<td>Felt and Lezoff, 1996 (85)</td>
<td>PND 10–24</td>
<td>2–4 mg/kg (1%)/PG10–G10, G0–21, G10/PND 10, PND 0–21</td>
<td>PNDs 24, 66, 8, 12, 16/47% (G 0–21, PND 0–21)</td>
<td>PNDs 75–90= 107% (G 0–21, PND 0–21)</td>
<td>Home orientation: smaller % to reach home: PND 6 (PG 10-G 10*, G 0–21*, G 10-PND 10*, PND 0–21*) Fewer no. of octants entered: PND 8, 12 (PG 10-G 10*, G 0–21*, G 10-PND 10*, PND 0–21*)</td>
<td>Brain iron</td>
</tr>
<tr>
<td>Felt et al, 2006 (185)</td>
<td>4 mg/kg (10%)/ED 5–PND 7</td>
<td>to PND 20/ED 5–PND 20</td>
<td>PNDs 24–110*</td>
<td>PND 10, 85%; PND 25, 75%; &gt;PND 34, NS</td>
<td>Sensorimotor function (PNDs 24–90): vibrissa-stimulated forelimb placing*** (PNDs 24, 27), sticker test (PND 35, NS; PND 70***), naturalistic grooming* (PND 90); Activity: activity monitor (less rearing*) (PND 70); water maze (PND 35): more days to criteria*, longer path length***, greater % of path length close to wall***, Response to novelty (PND 35–40), NS; spatial alternation (PND 70–110), NS; Morris water maze: greater swim distance: G 0–21, G 10-PND 10* Activity wheel: NS</td>
<td>Brain iron, monoamines, dopamine D2 receptor, dopamine and serotonin transporters</td>
</tr>
<tr>
<td>Glover and Jacobs, 1972 (176)</td>
<td>(a) iron free; (b) partially replenished/post-weaning*</td>
<td>d/Wistar rats</td>
<td>NR/34% (a); 82% (b)</td>
<td>85%</td>
<td>Activity: less activity*</td>
<td>Spontaneous activity: less spontaneous activity: 37 mo* Continuous performance attentional test: smaller % of sessions initiated: 34 mo*; greater % of correct rejections: 34 mo* Delayed nonmatch-to-sample test*: longer latency to choice: 34 mo* +G: reduced spontaneous activity (Fig 10)<strong>; lower inhibitory response to novel environment (Fig 6)*; more weekly behavioral changes (Fig 4)</strong>; less fearful (Fig 9)**; novelty preference test, NS; +PND: poorer performance on an object permanence task (Fig 5)<em>; more tense</em>; novelty preference test, NS</td>
</tr>
<tr>
<td>Golub et al, 1999, (186), 2000 (187)</td>
<td>10 mg/kg (10%)/30–37 mo of age*</td>
<td>/tha/ta monkeys</td>
<td>30 mo/103%; 34 mo/103%; 37 mo/97%</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>Hunt et al, 1994 (179)</td>
<td>ID+A: 4.6 mg/kg (4.2%), ID-A: 19.9 mg/kg (18%)/≈PND 28–84</td>
<td>d/♀/SD</td>
<td>PNDs 42, 84* (♂), 32%, (♀) 95%</td>
<td>ID+A (86%); ID-A (NS)**</td>
<td>Activity (10 outcome measures): ID+A, less activity (all measures but movement speed); ID-A, dark cycle: less distance traveled*, more stereotypy time*; light cycle: more distance traveled*, less vertical movement time*, less stereotypy time*</td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>References</th>
<th>Diet (% of control)/duration of iron deficiency</th>
<th>Sex/strain</th>
<th>At time of testing/age/iron status (% of control)</th>
<th>Body weight reduction (% of control)</th>
<th>Experiment results for iron-deficient groups compared with controls</th>
<th>Other brain-related experimental measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kwik-Uribe et al, 2000 (177)</td>
<td>14 mg/kg (19%)−63G−PND 21/−63G-PND 75</td>
<td>δ ♀/SW mice PNDs 30, 45, 60/21%; PND 21/−63G−PND 75</td>
<td>Weights significantly less (♀; PND 28−49; δ; PND 28−70); body weight used as a covariate</td>
<td>Grip strength**: lower forelimb grip strength; −63G−PND 21/−63G−PND 75; lower hindlimb grip strength −63G−PND 21/−63G−PND 75; −63G−PND 30/−63G−PND 75 (♀, PND30)<strong>; Auditory startle response: lower startle response; −63G−PND 21 (♀; PND 30, 45, 1 of 5 trial blocks)<em>; −63G-PND 75 (PND 30, 1 of 5 trial blocks)</em>; Morris water maze: escape latencies (NOVA, NS); probe trial, NS</strong></td>
<td>Brain iron</td>
<td></td>
</tr>
<tr>
<td>Kwik-Uribe et al, 2000 (177)</td>
<td>12.5 mg/kg (17%)−56G−PND 60</td>
<td>δ ♀/SW mice PNDs 30, 40, 50/60% (hematocrit)**</td>
<td>PND 21; 74% (♂); 82% (♀); PND 60; 83% (♂); 93% (♀)</td>
<td>Neurobehavioral battery; less forelimb grip strength**: ♂<em>, ♀</em>; Δ (♀, PNDs 30, 50)<em>; less hindlimb grip strength**: ♂</em>, ♀*; Δ (♀, PNDs 30, 40)</td>
<td>Brain iron, lipid oxidation (TBARS)</td>
<td></td>
</tr>
<tr>
<td>Massaro and Widmaier, 1981 (180)</td>
<td>7 mg/kg (19%)−40**</td>
<td>♂/Phases 2 and 3 (♀=60%)</td>
<td>Day 1: 91%; day 40: NS</td>
<td>Transfer learning test: deficits in ability to transfer a learned association from a visual to an auditory stimulus**</td>
<td>Brain iron</td>
<td></td>
</tr>
<tr>
<td>Pinero et al, 2001 (168)</td>
<td>3 mg/kg (7.7%)/−PND 10−35, PND 10−21, PND 21−35</td>
<td>♂/♀/SD PNDs 14, 17, 20, 27, 34/−PND 21: 69% (♀; PND 21−35); PND 35: 38%; (PND 21−35, 24% (PND 10−35)</td>
<td>PND 10−21; 79% at PND 21, 91% at PND 35; PND 21−35: 82% at PND 35; PND 10−35: 54% at PND 35</td>
<td>Activity: shorter distance traveled: PND 10−35 (PND 20, 27, 34)<strong>; PND 21−35 (PND 34)</strong>; PND 10−21 (PND20), PND 34**: slower rate of habituation: PND 10−21 (PND 20); PND 10−35 (PND 34)<strong>; fewer repeated movements: PND10−35 (PND 27, 34)</strong>; PND21−35 (PND14)**</td>
<td>Brain iron</td>
<td></td>
</tr>
<tr>
<td>Weinberg et al, 1980 (169)</td>
<td>iron deleted/PND 1−28</td>
<td>♂ ♀/rats PNDs 26−28/32% (hematocrit)</td>
<td>70−71% (♀); 70−71% (♂)</td>
<td>Open field: more rearing; less freezing*; Hole board: more ambulation on first test day*; head-dipping, NS</td>
<td>Brain iron, cytochrome oxidase*</td>
<td></td>
</tr>
<tr>
<td>Weinberg et al, 1979 (170)</td>
<td>iron deleted/PND 1−28</td>
<td>♂ ♀/SD PNDs 65−70/NS</td>
<td>≈70% (♀ PND 28); NS (♂ PND 60)</td>
<td>Hole board: NS</td>
<td>Brain iron</td>
<td></td>
</tr>
<tr>
<td>Williamson and Ng, 1980 (181)</td>
<td>2−3 mg/kg (=3.5%)/PND 90.174</td>
<td>♂ ♀/Wistar PNDs 104, 146, 174**/≈86%</td>
<td>NR</td>
<td>Taste aversion memory test**: poorer recall with immediate conditioning: PND 104**; better recall with delayed conditioning: PND 146**, PND 174**</td>
<td>Brain iron</td>
<td></td>
</tr>
<tr>
<td>Williamson and Ng, 1980 (182)</td>
<td>2−3 mg/kg (=3.5%)/PND 90.174</td>
<td>♂ ♀/Wistar PNDs 104, 146, 174/PND 174 (NS)</td>
<td>NR</td>
<td>T-maze: PND 104: fewer trials to criterion*, fewer errors**: PND 146: longer time to complete the maze**; PND 174; less time to complete the maze*</td>
<td>Brain iron</td>
<td></td>
</tr>
<tr>
<td>Yehuda et al, 1991 (171)</td>
<td>5.3 mg/kg (2.7%)/PND 42−69</td>
<td>♂ ♀/SD PND 69/NR; after repletion: ≈PND 90.9%/NR</td>
<td>Significant weight differences (PND 70, 91%); body weight used as a covariate</td>
<td>Water Y-maze ± electric shock: −electric shock. More trials to reach criterion (PND 69*, 90**; more errors (PND 69**); +electric shock: fewer errors (PND 69**)</td>
<td>Brain iron</td>
<td></td>
</tr>
<tr>
<td>Yehuda et al, 1986 (172)</td>
<td>Iron free PND 42−69</td>
<td>♂ ♀/SD PNDs 42, 49, 56, 63, 70/67% (PND 56), 58% (PND 63), 40% (PND 70); after repletion: 63, 70, 77, 84, 91%</td>
<td>82% (♀ PND 70)</td>
<td>Water Y-maze: more trials to reach criterion (before** and after*** repletion), errors**, and longer escape latencies** (PNDs 49, 56, 63, 70); duration of iron deficiency***</td>
<td>Brain iron</td>
<td></td>
</tr>
</tbody>
</table>
| Yoodim et al, 1981 (173)          | Iron free/28-d period (age NS; wt = 100−150 g) | ♂ ♀/SD PNDs 42.47%−42% | Beginning of experiment, NS | Activity ± Δ-daamphetamine: light: increased activity; dark: reduced activity (Δ-amphetamine*, +amphetamine**); Stereotypic behavior ± apomorphine: light: more stereotypic behavior***; dark: less stereotypic behavior*: after 1 wk repletion: NS** | Brain iron, body temperature (Continued)
suggest the need for caution when interpreting the results.

5-HT, 5-hydroxytryptamine; L-DOPA, levatory 3,4-dihydroxy-L-phenylalanine.

References

TABLE 2 (Continued)

<table>
<thead>
<tr>
<th>References</th>
<th>Diet (% of control)/duration of iron deficiency</th>
<th>Sex/strain</th>
<th>At time of testing (age/iron status (% of control))a</th>
<th>Body weight reduction (% of control)b</th>
<th>Experimental results for iron-deficient groups compared with controlsa</th>
<th>Other brain-related experimental measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Youdim et al, 1980 (174); Youdim and Green, 1977 (175)</td>
<td>Low iron/35-d period (age NS; wt = 80 g)</td>
<td>d/SD</td>
<td>44/13%</td>
<td>NR</td>
<td>Activity ± (a) + l-tryptophan; (b) + L-DOPA; (c) + 5-methoxy-N,N-dimethyl-tryptamine; (d) + methamphetamine or apomorphine: after period of iron deficiency: less hyperactivity (a**, b*, c**, d**); after 8-d period of repletion: NS (a, b)</td>
<td>5-HT, dopamine, noradrenaline, tryptophan, hydroxylase, aldehyde dehydrogenase, adenylyl cyclase, protein synthesis</td>
</tr>
</tbody>
</table>

1 All experiments used rodents unless otherwise specified: PG, prior to mating; G, gestation; PND, postnatal day; SD, Sprague-Dawley rats; SW, Swiss-Webster mice; ED, embryonic day; 5-HT, 5-hydroxytryptamine; L-DOPA, levatory 3,4-dihydroxy-L-phenylalanine.

2 Iron status based on hemoglobin concentration unless otherwise indicated; the only values shown are those that were significantly different from controls. Values are at the time of testing unless otherwise indicated.

3 Values are at the time of testing unless otherwise specified.

4 *P < 0.05, **P < 0.01, ***P < 0.001. When an interaction involving diet was reported, the P value is assumed to be the procedure-wise error rate used for post hoc tests. When an interaction involving diet was reported but post hoc test results were not reported, the main effect of diet is presented in parentheses. Groups are shown by period of iron deficiency. When multiple tests were conducted, results for tests that did not show statistically significant differences between iron-deficient and control animals are not shown.

5 No P value provided.

6 Estimated from Figure 1 of the cited reference.

7 Quantitative results not provided.

8 Results not shown. No significant group differences for escape latency or swimming velocity; statistically significant differences observed only for individual trials.

9 Investigators state that after 2 wk of an iron-sufficient diet (ie, at PND 34), hematology did not differ significantly from that of controls. Values are not provided for PNDs before PND 34.

10 Authors indicate that after 2 wk of the iron-sufficient diet, the weight of experimental and control animals did not differ significantly; data are provided only for weights at PND 120.

11 Duration not reported.

12 No statistical analysis was conducted, but group differences were large. Diurnal pattern appeared to be reversed in both ID+ A and ID-A groups compared with controls.

13 One-half of the iron-restricted group was supplemented with iron from 34 to 37 mo of age. Test results for this group at 37 mo of age were not significantly different from those of the controls or iron-restricted unsupplemented groups. Diets of iron-restricted groups were also restricted for zinc.

14 The delayed nonmatch-to-sample test is part of the Cambridge Neuropsychological Test Automated Batteries test program.

15 Investigators indicate that dams were ID+ A in the third trimester and infants were born with compromised iron and hematologic status; at 1 and 4 mo of age, hematologic status was normal.

16 The large number of statistical tests in this study, the fact that all but one result considered positive were only weakly significant (P < 0.05), and the apparent lack of a consistent pattern of effect suggest the need for caution when interpreting the results.

17 The text indicates P = 0.04; the legend in Figure 4 indicates P = 0.01.

18 Continuous activity monitor recorded 10 activity measurements.

19 It is indicated that body weights are “final body weights.”

20 P values indicate differences between the control group and each iron-deficient group in regression models constructed for each activity measure over the 8-wk period of the experiment.

21 9 wk before mating.

22 The same animals were tested on all 3 d.

23 Data presented in a figure; differences appear to be ~10–15%.

24 Statistical significance indicated was after adjustment for body weight differences by including body weight as a covariate in the analysis.

25 Animals tested in the Morris maze were naive.

26 Results of significant post hoc tests conducted after a nonsignificant ANOVA are uncertain.

27 Dams were also subjected to tail bleeds.

28 Iron status not reported on other PNDs.

29 Results were adjusted for effects of weight differences between iron-deficient and control offspring.

30 Age not specified; animals weighed ~90 g at the beginning of the study period.

31 Habituation (the first of 3 phases of the experiment) took place during the first 28 d; the last 2 phases occupied the remainder of the 40-d experimental period.

32 Offspring used in the 3 tests (open field, hole board, passive shock avoidance) were naive but from the same litters.

33 Results for these endpoints were obtained in a separate experiment.

34 Animals tested at each time point were naive.

35 Activity levels are also reported, but methods are not described.

36 After initial testing, each group was fed the control diet (200 mg Fe/kg) for 3 wk and then retested.

37 Actual weights not reported.

38 Separate groups of rats were tested at each time period.

39 After initial testing, each group was fed the control diet (220 mg Fe/kg) for 3 wk and then retested.

40 Trials to reach criterion increase with duration of iron deficiency.

41 End of the period of iron deficiency and after a 1-wk repletion period.

42 End of 28-d period of iron deficiency.

43 Data not shown.

44 End of period of iron deficiency and after an 8-d repletion period.
uncertain. In a recent experiment from this group (188), the hematologic status of the offspring was normal at the time of testing, but, at birth, the dams and infants of the prenatally iron-restricted group were ID+A. The postnatally restricted group had lower ferritin concentrations than did the controls but evidently did not develop anemia.

**Dose-response analysis**

The only study that compared groups with different degrees of iron restriction (179) used an activity monitor to compare 10 outcome measures over a period of ≈6 wk. Detailed results were reported for distance traveled, movement speed, vertical movement, and time spent in repetitive movements. Whereas there were statistically significant differences in most outcome measures for both severe and marginally iron-restricted groups and controls, as shown in Table 2, the pattern of effect was not consistent.

**Pre-compared with postweaning iron restriction**

As shown in Table 2, dietary intake of iron was restricted either during gestation or before weaning (85, 168–170, 177, 178, 184, 185) or after weaning (48, 168, 171–176, 179–182). In all cases, statistically significant performance differences between iron-restricted and control groups on at least some performance tests were reported.

**Test specificity**

Twelve of the studies listed in Table 2, from 6 different laboratories, included performance tests that measured endpoints involving learning or memory (85, 169–172, 177, 180–182, 185–188), and 13 studies included tests that measured motor, exploratory or homing behavior, or physical measures of development such as grip strength and balance (48, 85, 168–170, 173, 174, 177–179, 181, 184–188). As shown, the most commonly measured endpoint was activity, and several independent groups reported decreased activity in iron-restricted animals (48, 168, 169, 174–176, 185–188).

The only test that involved cognitive endpoints that was used by more than one independent group (85, 177, 185) was the Morris water maze, a spatial learning test (81, 190). Two laboratories reported that iron-restricted groups performed less well than did controls. In rats, Felt and Lozoff (85) tested 10–12-wk-old rats and reported that iron-restricted groups performed less well in the Morris water maze, a spatial learning test (81, 190). In our opinion, although all investigators discussed possible interpretations that could explain their results, the body of data taken together does not present a consistent or clear picture.

**Repletion**

As shown in Table 2, in all studies that examined performance in iron-repleted animals that had been iron-restricted during gestation or before weaning, performance was significantly below that of controls in at least some tests (85, 168, 170, 177, 184–187). One research group examined performance after repletion of animals that had been iron-restricted after weaning or as adults (172–175). Although normal motor activity was restored after iron repletion (173–175), performance in a water Y-maze was not (172).

**ARE CAUSAL CRITERIA SATISFIED?**

There has been considerable discussion in the published literature of causation relevant to dietary iron deficiency and cognitive or behavioral performance; several examples are cited (1, 38, 84, 99, 109, 191). Most discussion has focused on effects in humans, particularly on difficulties in designing treatment trials due to ethical prohibitions against including placebo control groups, on potential confounders such as clinical effects of anemia that do not directly involve the brain, or on socioeconomic or nutritional disparities between comparison groups. Discussion of the 5 causal criteria below includes these and other considerations in brief summary and also reflects the importance of taking into account results from animal studies in making overall judgments about causality.

**Consistent associations**

As indicated in the body of the review, associations between ID+A and deficits in cognitive or behavioral performance in children are consistently observed (1, 37–39, 90, 91, 93–96, 142, 192, 193). In animals, though decreased motor activity is consistently associated with ID+A, performance in cognitive or related tests has been less widely studied, and independently replicated results are not yet available (see text and Table 2).

**Dose-response relations**

In humans there is limited evidence of dose-response effects. Performance deficits increased with increasing severity of ID+A in experiments that used either the Bayley scales in children <2 y old (102) or intelligence quotient and school tests in older children (194). A significant correlation was also observed between the severity of anemia and conduction time in a test measuring auditory evoked potentials (129). In addition, scores on performance tests by children with ID+A were consistently lower than those of children with ID-A in most studies that directly compared both groups (102, 120, 134, 136, 144). It is also noted that, although not clearly relevant to the effects of iron deficiency on the brain, a dose-response relation between varying degrees of ID+A and work performance has consistently been observed (49, 112–116).
The rodent model provides an excellent opportunity to determine the shape and character of dose-response relations and to establish threshold levels of iron deficiency above which effects might not be expected to occur. This information is essential to determine which effects of iron deficiency might be expected under less-severe conditions. However, surprisingly few dose-response studies have been carried out (147, 179, 195, 196), only 2 of which included measurements related to possible effects on the brain (179, 196). One of these studies (196) observed a decrease in brain succinate dehydrogenase and cytochrome oxidase activities when animals were fed diets containing 2 mg Fe/kg but not when diets contained 6 mg Fe/kg, both of which resulted in severe ID+A. In the other experiment (179), motor activity was altered in both the ID+A and ID-A animals (see Table 2 and Discussion in the text). Clearly, more work is needed.

Ability to manipulate the effect

As discussed above, in most cases, performance deficits in children with ID+A >2 y of age were ameliorated by iron treatment, thus providing important evidence of ability to manipulate the effect (104, 118, 119). This finding is supported by the reversibility by iron treatment of motor activity (173–175), brain iron concentration (183, 197), and dopamine D2 receptor density (183, 198) in rodents when the period of iron restriction occurred after weaning.

In contrast, performance deficits were generally more difficult to reverse in children <2 y of age (102, 103, 117). Several possible explanations for this observation have been discussed (117), as indicated in the text. However, it is of interest that similar observations were consistently made in the rodent model, where groups differed only in dietary iron availability. In rodents, when the period of iron restriction occurred during gestation or before weaning, deficits in motor activity (85, 87, 168, 170, 177, 186, 187, 199), brain iron concentration (85, 170, 183, 200–202), and D2 receptor density (86, 183, 203) were all more resistant to reversal by iron treatment than when restriction occurred after weaning. As in the human studies discussed above, these observations are compatible with more severe effects of iron deficiency occurring during early development.

Specificity of cause and effect

As pointed out by many reviewers (38, 84, 91, 191), the specificity of cause in human observational and treatment trials involving children with ID+A is potentially confounded by nutritional and socioeconomic factors that could theoretically have influenced outcomes. Preventive trials provide the opportunity to better establish specificity of cause. However, as indicated in the text, results of the small number of preventive trials that have been conducted are not definitive. Rodent studies more easily satisfy this criterion because the only difference between controls and test animals is the amount of dietary iron available.

The criterion of specificity of effect is difficult to satisfy in both systems. Even relatively mild iron deficiency has been reported to affect general energy levels in both humans (204, 205) and rodents (148, 149), which could affect performance on activity as well as many cognitive and behavioral tests. In addition, because the degree of iron restriction in most rodent studies is severe, potential confounding due to undernutrition is also a potential confounding factor (146, 150, 151, 206). Rodent experiments in which deficits in performance and other brain effects occurred in the same animals, such as decreases in brain iron (48, 85, 168, 170, 173, 177, 178, 181) and dopamine D2 receptor density (48, 174), provide evidence consistent with a specificity of effect, as do experiments that used weight controls (169–171, 183) or included weight as a covariate (177, 178).

Plausible biological mechanisms

Slower myelination and decreased activity of associated enzymes (50–52, 207), reduction in the density and affinity of dopamine D2 receptors, and other effects on monoamine neurotransmitter systems (47, 153, 167, 184, 185, 203, 208, 209) are associated with iron restriction in rodents and have been widely discussed as possibly responsible for the observed effects on motor and cognitive or behavioral performance (1, 49, 152, 153, 155, 165). Other biochemical and morphologic changes in the brain have also been observed after severe iron restriction in rodents, including decreased activity or concentration of proteins involved in energy metabolism (cytochrome C oxidase and cytochrome c) (146, 163, 164, 210, 211), slower dendritic growth (201), and alterations in neural metabolites in the hippocampus (202). Recently, Yehuda and Mostofsky (212) suggested that the primary effects of iron restriction might be on signaling molecules generated peripheral to the central nervous system and that these might in turn influence dopamine D2 receptor density. The concentration of heme is reduced in cells grown in vitro under conditions of iron restriction (164). Atamna et al (213) showed that heme deficiency causes the mitochondria to release oxidants, which could jeopardize a variety of cellular functions in the brain. The reader is referred to additional expert reviews for general discussion of possible mechanisms (1, 49, 97, 153–156, 161, 166).

DO COGNITIVE OR BEHAVIORAL DEFICITS RESULT FROM IRON DEFICIENCY WITHOUT ANEMIA?

Expert reviewers suggest that this question is unresolved (38, 47, 49, 90, 109, 134, 163). However, despite an unfortunately small number of studies involving ID-A, some evidence, both direct and indirect, does suggest that cognitive or behavioral performance and brain function could be affected, particularly if deficiency occurs during critical stages of development. As discussed in the text, direct evidence from a relatively small number of human studies is suggestive but limited (Table 1).

Indirect evidence cited by expert reviewers in discussions of ID-A include I) sensitivity of the brain to iron deficiency during early development (1, 38, 46, 47, 49, 90, 146, 156, 214), 2) variable sensitivity to iron deficiency of different brain regions and enzymes (1, 46, 47, 49, 163), 3) effects on some enzymes before anemia in iron-restricted rodents (147), and 4) prioritization of tissue iron to red blood cell formation during development without a decrease in hemoglobin, such as occurs after fetal hypoxia (27, 97, 163).

Two additional points should be noted. First, the relation between dietary iron deficiency and functional brain iron is essentially unknown because such a high percentage of brain iron is stored in ferritin or is associated with hemosiderin or transferrin (47, 58, 215, 216) that assays are not sensitive enough to detect relatively small changes in the concentration of functional iron. Thus, the kinetics of region-specific decreases in functional iron and the regional availability of stored iron for functional needs are unknown. Second, in rodents, effects on brain function occurred when brain iron (storage iron) was <30% depleted (169,
IRON AND COGNITION 941

170, 177, 178, 183, 199, 217–219), which could suggest either that functional deficits occurred before the depletion of iron stores in the brain or that a significant percentage of stored brain iron was not available to all brain regions and functions under iron-restrictive conditions. Of relevance are observations that most iron stored in the brain is located in only a few brain regions (220), and no correlation was observed between loss of brain iron and cytochrome oxidase activity when values were determined across individual brain regions (163, 221).

Thus, although direct evidence demonstrating an effect of ID-A on cognitive, behavioral, or other brain functions is limited, until shown otherwise, as some have suggested (97, 163), it seems prudent to assume that a gradation of effects of iron deficiency occurs in the brain, with milder anemia and and ID-A resulting in perhaps more subtle but still potentially adverse brain effects, particularly if they occur during sensitive periods of development.

RESEARCH NEEDS

Areas in which further research is needed have been discussed by investigators: 1) critical periods of development that are particularly sensitive to iron deficiency (1, 91, 161, 184); 2) regional differences in brain sensitivity to iron deficiency (1, 160, 184); 3) effects in high-risk persons, such as low-birth-weight or premature infants (27, 38); 4) effects of maternal iron deficiency during the prenatal period on pregnancy outcome (e.g., prematurity; 222) and maternal nurturing (223, 224); 5) effects in persons with multiple deficiencies (91); 6) uncertainties in extrapolating results from animals to humans (49); and 7) effects in longitudinal and long-term follow-up studies (38, 84) and in studies of sufficient statistical power to detect weak but possibly meaningful effects (37, 38).

Expert reviewers have also commented on the need to develop more sensitive and specific cognitive tests (1, 44, 84, 162, 163). In this regard, we note that there are many well-established and widely used tests in rodents that target various aspects of learning, memory, attention, and emotion. For a general review, see the article by Whishaw and Kolb (225) and our recent reviews (5, 6), which include references to methods for a wide variety of performance tests. It is striking that so few of these tests have been used to examine effects in iron-restricted animals.

The almost complete absence of dose-response analysis in this field is striking. Rodents are an ideal system for such studies, which are needed to understand the range of iron-restrictive conditions over which effects occur. In addition, it is clear that much is still to be learned about the relation of storage iron and functional iron in the brain, particularly on a regional basis. We also note the dearth of studies that have directly examined brain iron in humans during development and under conditions of iron deficiency. Noninvasive techniques are now available that could be applied to discern developmental patterns of iron accumulation in the iron-deficient developing human brain (226).

CONCLUSIONS

Although most causal criteria are supported by at least some evidence from either human or animal studies, significant gaps suggest that it would be premature to conclude that a causal connection exists between iron deficiency per se during development and subsequent cognitive or behavioral performance. In humans, although several causal criteria have been at least partially satisfied, the specificity of both cause and effect has not been clearly established. Animal studies provide important support for results in humans and also supply information that cannot be obtained in human studies. In the rodent experiments discussed here, plausible biological rationales have been identified, the specificity of cause has been better established than in humans, an association between deficits in motor activity and severe ID+ A have consistently been observed, and the reversibility or irreversibility of effects of iron treatment parallel observations in humans. However, the specificity of effect has not been established in animal studies, relatively few cognitive and behavioral tests have been conducted, results have not been independently replicated, and dose-response relations have essentially not been investigated.

Although there is some logical support for and suggestive evidence from a few human studies of the possible effects of ID-A on cognitive or behavioral development, the literature is surprisingly uninformative on this extremely important topic, primarily because so few studies have been conducted. Because 2 billion women and children are iron deficient worldwide, and the greatest prevalence of ID± A in the United States is among adolescent girls (9–16%) and children during the brain growth spurt (7%) (16, 18, 20, 21, 44), further studies are clearly needed. It should be noted that results are just beginning to appear from a multilaboratory cross-species research effort aimed at addressing some of the issues discussed in this review (184, 188).

We are very grateful for the helpful comments on various drafts of this manuscript from Fernando Viteri, Hani Atamna, John Beard, Ash Lal, Elizabeth Theil, and Patricia Wainwright. We also greatly appreciated the thoughtful comments of the AJCN reviewers. We thank Teresa Klask for administrative assistance and Priscilla Vargas for help with reference acquisition.

JCM searched and analyzed the literature and wrote the review. BNA suggested the topic, gave advice, and edited the review. Neither of the authors had a personal or financial conflict of interest.

REFERENCES

10. Siega-Riz AM, Popkin BM. Dietary trends among low socioeconomic


62. Siimes MA, Addiego JE Jr, Dallman PR. Ferritin in serum: diagnosis of...
116. Li R, Chen X, Yan H, Deurenberg P, Garby L, Hautvast JG. Functional...


189. Hill JM, Switzer RC III. The regional distribution and cellular local-