Effects of discontinuing coffee intake on iron deficient Guatemalan toddlers' cognitive development and sleep

P.L. Engle[a,*], T. VasDias[b], I. Howard[c], M.E. Romero-Abal[d], J. Quan de Serrano[d], J. Bulux[d], N.W. Solomons[d], K.G. Dewey[e]

[a]*Department of Psychology, Cal Poly State University, San Luis Obispo, CA 93407, USA*
[b]*University of Hawaii, Honolulu, HI, USA*
[c]*Instituto Guatemalteco de Seguridad Social, Guatemala City, Guatemala*
[d]*Center for Studies of Sensory Impairment, Aging, and Metabolism, Guatemala City, Guatemala*
[e]*Department of Nutrition, UC Davis, Davis, CA, USA*

Abstract

Coffee is commonly given daily to toddlers in Guatemala. Possible negative effects of coffee ingestion on cognitive development and sleep patterns were assessed in 132 children 12–24 months of age who had received coffee for >2 months and were iron deficient on at least one indicator. Children were stratified by initial hemoglobin (A = anemic, Hgb <10.5 g/dl; NA = 'non-anemic', Hgb ≥ 10.5 g/dl) and were randomly assigned to an experimental group (S = substitute consisting of sugar and coloring), and a control group (C = continuation of coffee) (42 C-NA; 53 S-NA; 18 C-A; and 19 S-A). Anemic children were provided Fe supplements for 2–3 months. Compliance was assessed every 2 weeks. After 5 months, testers masked to treatment group and anemia evaluated children with the Bayley Scales of Infant Development II in a central location. Scores were the Mental Development Index (MDI), the Psychomotor Development Index (PDI), and scales from the Behavior Rating Scale (BRS). The child's sleep in the previous 24 h was assessed with a set of standardized sleep questions to the care giver on the first visit and every 2 weeks thereafter. No significant effects of treatment on test scores or BRS ratings were found. In the 24 h period reported on at the final visit, children in the Substitute group slept more during the night and overall (night plus naps) than children in the Coffee group, a difference not found at the first visit. No differences were

found in sleep difficulty or number of times waking at night. Women's reported coffee intake per day during pregnancy was associated with lower BRS ratings, even after controlling for SES and child age. The effects of postnatal coffee ingestion in Guatemala were seen for sleep duration, but not for cognitive development. Prenatal coffee ingestion was negatively associated with Behavior Rating Scales and should be investigated further.

Keywords: Bayley Scales of Infant Development II; Caffeine; Coffee; Guatemala; Iron status; Sleep

1. Introduction

Coffee is a common beverage not only among adults but also among children in Guatemala, the sixth largest producer of coffee [1]. It is one of the first liquids given to infants after breast milk [2,3]. Coffee is considered a beverage only for adults in the United States, but it is served daily to Guatemalan children as young as 6 to 8 months of age [4]. To date, little is known about the effects on children of coffee intake.

There are at least two mechanisms by which coffee intake could affect child behavior. Coffee might have a direct effect on sleep and cognitive function through increasing excitability or inattention. It may influence cognitive function indirectly by impairing iron status, which could then be associated with slower cognitive development. Although this effect could be mediated by caffeine, caffeine is not the only metabolically active agent in coffee, and other constituents are thought to be responsible for the influence of coffee on iron status [5].

There is some evidence that coffee and/or caffeine intake affects behavior in both animals and humans. Postnatal coffee ingestion has been shown to influence brain function and brain chemistry in rodents [6,7]. Caffeine and the xanthines in coffee have been associated with greater reactivity and timidity in rodents and in humans [8], but these effects are much clearer in studies with rodents, when much higher dosages are administered [9]. A few experimental studies of caffeine intake by school-aged children, usually from sources other than coffee, suggest that moderate levels (e.g., 3 mg/kg/day) have relatively minor effects, whereas larger amounts (e.g., 10 mg/kg/day) can result in increased restlessness, excitability, and inattention [10–13]. For a child weighing 30 kg, 10/mg/kg/day would be equivalent to three cups of brewed coffee, or almost five cups of instant coffee. Data from adult subjects show similar results. At smaller doses of coffee, vigilance increases and memory appears better [14,15]. Under increased dosages (to 300 to 400 mg/day, about four cups of coffee, 6 mg/kg/day), functioning begins to decline, with increased reactivity, emotionality, and irritability [16]. Coffee appears to increase energy expenditure in adults with dosages as small as 1.5 mg/kg/day [17,18].

The hypothesis that there is an indirect effect of coffee on cognitive development has not been tested previously. However, it has been demonstrated that coffee interferes with iron absorption [19,20] and affects the iron status of pregnant women

and their infants [5]. Also, numerous studies have demonstrated an association between iron deficiency anemia and development, even when socioeconomic factors are controlled [21–27]. Thus it is reasonable to hypothesize that changes in coffee intake could affect iron status and in turn influence cognitive development. However, these affects could vary significantly as a function of the initial iron status of the child [28].

A number of studies have shown that coffee/caffeine interferes with sleep onset in adults. Bonnet and Arand [18] found that doses of 400 mg of caffeine taken three times a day (about 16 mg/kg/day) reduced the subjects' sleep efficiency to levels associated with insomnia. A standard cup of brewed coffee contains on average 98 mg of caffeine (range 40–180), whereas instant coffee contains 65 mg of caffeine (range 30–120) and decaffeinated coffee contains only 3 mg of caffeine (range 1–5) [29]. Doses of 300 to 400 mg of caffeine prior to sleep onset resulted in 30 to 80 min reductions in total sleep time [30,31]. Smaller doses also have disruptive effects on sleep. A dose of 100 mg resulted in a delay in sleep onset that was followed by poor quality sleep for the next 3–4 h [32]. One study reported that the alerting effect of coffee lasted between 5.5 and 7.5 h in two different experimental groups [33], which is consistent with the 3 to 5 h half-life of caffeine among adults. Some data suggest that the half-life of caffeine is much longer for pregnant women and for newborns [29]; in infants it decreases gradually with age during the first year to the adults' norm [34]. No studies examining the effects of caffeine on sleep patterns of infants and toddlers were found in the literature.

The present study was designed to test the effects of discontinuing coffee among Guatemalan 12–24 month olds using a randomized intervention trial. Children were initially classified on iron status, and divided into anemic and 'non-anemic' groups who were then randomly assigned within group. Sample size estimates were based on the power necessary to detect an effect size of 0.50, or a DQ of 7.5 points, which might have been expected based on other investigations of iron status.

In the study, coffee rather than caffeine was the independent variable, and there was no attempt to isolate the effects of caffeine from the coffee. Measurements were taken on iron status, growth, morbidity, daily sleep patterns, and cognitive development, as well as compliance with the intervention and food and fluid intake during the course of intervention. The study had three additional dependent measures assessed at the beginning and end of the trial: iron status of the child, morbidity, and growth. Results for these three dependent measures have been reported elsewhere [35,36]. This report summarizes the effects of the intervention on children's cognitive development, behavioral responses to the testing situation, and sleep patterns.

2. Methods

2.1. Study design

Several pilot studies were conducted to develop a coffee substitute for Guatemalan toddlers and to assess its acceptability. We then initiated a randomized, intervention trial with two groups: (1) the Coffee group, which continued to drink coffee during

the 5-month intervention period, and (2) the intervention group (the Substitute group), which was provided with a substitute for coffee consisting of an 'instant' brown-colored sugar cube to be mixed with hot water in the home. The amount of sugar that this provided (approximately 6 g/100 ml) was similar to the average amount used in preparing coffee for children. We chose this substitute, rather than more nutritious beverages such as juice or milk, in order to test specifically the impact of coffee independent of alterations in nutrient intake. In the acceptability trials, mothers reported that their children liked the substitute as much as, or more than, coffee. Mothers in the randomized trial were not told the actual purpose of the study, but to be eligible to participate they had to be willing to stop giving coffee to their children if assigned to the intervention group.

Because the supplement had to be delivered regularly to the Coffee group, there was no way that the home interviewers could be unaware of the child's group. However, the interviewers were well educated and trained to reduce the chances of bias. Testing of the Bayley Scales of Infant Development was masked, since the psychometrists administered the test at a central site, and they were not informed of the children's status. The Bayley Scales were administered within 15 days of the last home visit, so that the Substitute group could continue to receive it up through the date of testing. The trial was considered to be completed when the test was administered and the final blood draw occurred. The institutional review board at the University of California, Davis and the Human Subjects Committee of the Center for Studies of Sensory Impairment, Aging and Metabolism in Guatemala approved the study protocol.

2.2. Subjects

Subjects were recruited through door-to-door canvassing in dispersed low-income neighborhoods of Guatemala City. At the time of recruitment, children were screened for eligibility, a blood sample was taken to determine iron status, and anthropometric measures were completed. Criteria for inclusion were that the child was between 12 and 24 months of age, had received coffee for at least 2 months, had an intake of more than 90 ml/day of coffee, and was iron deficient (hemoglobin <11.5 g/dl, hematocrit $<35\%$, and/or zinc protoporphyrin-heme ratio (ZPP/H) >80 μmol/mol heme). Children with chronic medical conditions were not included in the study.

Subjects were categorized into those who were initially anemic (Hgb <10.5 g/dl) or 'non-anemic' (Hgb ≥ 10.5 g/dl). The children were then randomly assigned to the Substitute or the Coffee groups. The mothers of the anemic children were provided iron supplements to give to the children (5 mg/kg/day) for 2 to 3 months.

2.3. Instruments and procedure

Eligible subjects were visited in the home shortly after recruitment to obtain demographic data, evaluate dietary intake using a food frequency questionnaire, and assess the child's sleep patterns during the previous night and day. At this visit, subjects in the intervention group were given their first 2-week supply of the

substitute and instructed in its use. Every 2 weeks, all subjects were visited in the home to assess the child's intake of fluids, compliance with using the substitute (Substitute group) or continuation of coffee intake (Coffee group), morbidity since the previous visit, and the child's sleep during the previous night and day. At the completion of the 5-month intervention period, the final blood sample was collected, anthropometric and food-frequency assessments were repeated, and the Bayley Scales of Infant Development II (BSID II) were administered. Interviewers were all college educated, female, and had been trained by the nutrition team at Cessiam, which has a long history of biomedical and nutritional research in Guatemala.

2.3.1. Blood sampling and analysis

The most common measure of iron deficiency anemia is hemoglobin (Hgb), which assesses the adequacy of the supply of iron for red blood cell formation. Hematocrit (Hct) is a measure of the packed red cell volume. In a healthy adult male, about two-thirds of the body's iron is in hemoglobin, in the red blood cells [37]. A third measure used here is ZPP/H, reflecting the increase in protoporphyrin in red blood cells that occurs when insufficient iron is available to combine with the protoporphyrin to form heme. Thus, high values of ZPP/H reflect an insufficiency of iron [37].

2.3.2. Anthropometry

Each child's weight and length were measured at the beginning and end of the study. Weight was measured to the nearest gram using a digital platform balance. Two different interviewers measured recumbent length to the nearest centimeter using an infant length board. *z*-Scores for weight-for-age, length-for-age and weight-for-length were calculated using WHO/CDC reference data (Epi-Info Version 6.03).

2.3.3. Dietary and coffee intake

Dietary patterns were evaluated at the beginning and end of the study using a food-frequency questionnaire. During the intervention, data on intake of liquids were collected continuously at each home visit using a 2-week recall of frequency and approximate amount consumed for each item. At recruitment and at each home visit, detailed information was obtained on the type of coffee used in the household, the amount of water used in its preparation, and any additional items added (e.g., milk, sugar, more water) in preparing coffee for the child. This information was used to calculate the 'dose' of actual dry coffee per kilogram body weight of the child. In this calculation, the weight of instant coffee was multiplied by a factor of 2 relative to ground coffee (based on the fact that caffeine concentration per gram dry weight of instant coffee is about twice that of ground coffee). Because some brands of coffee in Guatemala contain substantial amounts of non-coffee fillers, such as roasted grains, samples of each brand were analyzed for caffeine content to indicate the 'strength' of the coffee. These data were used to calculate another index of coffee dose based on caffeine concentration, and were used to calculate initial caffeine intake per kilogram of child's weight per day. Attempts were made to assess coffee intake with saliva

samples, but children were very resistant to having the samples taken, and technical problems occurred in using the samples.

Mothers also reported the number of cups of coffee ingested during each trimester of the pregnancy with the subject child. Their daily caffeine ingestion per day was calculated based on the assumptions that the mother was drinking from a standard cup of 150 ml, and was drinking the same kind of coffee that she was giving to her child. A log transformation of the measure was used to insure normality.

2.3.4. Sleep patterns

At each home visit, the mother was asked to recall information about the child's sleep pattern the previous day. Specifically, she was asked about the number of naps and the length of each, the time the child went to sleep at night and the time the child awakened in the morning, difficulty in getting the child to fall asleep in the evening, and whether and how often the child awoke during the night. When and where the child took naps and time to sleep onset was assessed as well.

Five variables were constructed for each visit from the standardized sleep assessment questionnaires: total time slept at night, total time slept during the day, total time slept (day and night), total number of times the child awoke at night, and average rating of the child's difficulty going to sleep on a four-point scale. Nap times were based on the mother's estimates of duration, but sleep at night was calculated based on the time the child was put to bed, and the time the child awoke in the morning. In pretesting, it appeared difficult for mothers to recall how long their children were awake at night, and numbers were usually rounded (10 min, 30 min) so the estimate of number of times the child awoke during the night was used instead.

2.3.5. Cognitive development

The Bayley Scales of Infant Development II (BSID-II) is an individually administered battery of test items which assess the current developmental functioning of infants and children. The examination consists of three scales: the Mental Development Index (MDI), the Psychomotor Development Index (PDI), and the Behavior Rating Scale (BRS). The MDI items assess memory, problem solving, habituation, generalization, classification, early number concepts, vocalizations, language, and social skills. The PDI measures control of the gross and fine muscle groups. Gross motor movements include rolling, crawling, sitting, standing, walking, running, and jumping, and fine motor manipulations involve prehension, adaptive use of writing implements and imitation of hand movements. The BRS assesses qualitative aspects of the infant's test-taking behavior and includes three sub-scales: (1) Orientation and Engagement, (2) Emotional Regulation, and (3) Quality of Movement [38].

The new version of the Bayley Scales of Infant Development, Version II, was published just before the initiation of the study. No normative data were available for Spanish-speaking children. The ethnic distribution of the normative sample was representative of the US population, therefore 11% of the sample was Hispanic [38]. However, only English-speaking Hispanics were tested, since the test had not been translated into Spanish for norming the instrument. For this study the test was adapted

to urban Guatemalan children's environment through an intensive series of workshops with two of the authors (PLE, US psychologist, and IH, a Guatemalan psychologist), and the three testers, all of whom were psychologists with a college degree. The instructions were translated into Spanish, tested on a number of children, and the materials were examined for culturally inappropriate items. The test was found to be appropriate for this urban environment, and materials did not have to be changed. In a few cases, vocabulary and specific instructions were changed.

The Behavior Rating Scales (BRS) is a totally revised version of the Infant Behavior Record (IBR) used in the original Bayley Scales. It is now based on factor scores of the items [38]. In the past, the IBR had been found useful for detecting subtle, qualitative differences such as those found with a clinical diagnosis [39], but the scoring system was not well developed. When scored properly, it was found to be more predictive of later functioning than the MDI and PDI scales in some studies [40]. No data are yet available on the predictive validity of the new BRS.

Each of the 30 items on the BRS is a five-point Likert scale, and the tester has to rate the child's behavior in the testing situation. Each item is anchored with a specific description. For example, item 16 measures 'enthusiasm toward tasks', one of the items in the Orientation and Engagement Factor, described as ''The degree to which the child exhibits deep concentration, coupled with excitement or delight, in the materials or tasks'' [38] (p. 185). The values for the item are (1) consistently unenthusiastic, no particular interest beyond attending to the tasks; (2) typically unenthusiastic, enthusiastic in one or two instances; (3) unenthusiastic half the time; (4) typically enthusiastic, unenthusiastic in one or two instances; and (5) consistently enthusiastic. The Orientation and Engagement Factor for this age group includes nine items which clustered together in factor analytic studies, and reflect the child's willingness to look at and engage both the tasks at hand and the examiner. The Emotional Regulation Factor includes 10 items reflecting the child's ability to adjust emotionally and adapt to the challenges of the testing situation, such as negative affect, ability to adapt to change in materials, attention, frustration, and cooperation. The total score combines items from the two subscales and the Motor Quality Factor scales (eight items).

The test was divided into five 'sets' as recommended in the manual, so that each child was administered a group of items depending on age. If the child performed above the ceiling on that set, a higher set was administered. If the child performed below the basal level, the next lower set was administered, in accordance with the instructions for the test. The total number of correct items for that set was converted into an index score based on the published age-based norms for the test. Corrections in the age of the child for prematurity were made only if the child was more than a month premature and was under 24 months. In the sample, 7% were more than a month premature, according to the mothers' reports. There was no difference in distribution of the subjects into the S and C groups by prematurity ($\chi^2 = 1.81$); however, premature children were more likely to be anemic ($\chi^2 = 11.78$, $P < 0.01$).

Three Guatemalan psychologists who were masked from the treatment group and hematological status of the children tested the children in a central location at the end of the study. Two of the three testers administered the test to each child; one person

asked the questions, and the other supplied materials and coded the child's responses. The care giver and child were brought to the testing site, and the care giver was present during the testing. The child's performance was evaluated on the 30 BRS items by one of the two testers at the close of the testing session, before the next child was tested. Two to three children were tested per day.

All three testers were trained to satisfactory levels of reliability ($>85\%$) on the test before initiating the study, and particular attention was paid to the reliability of the BRS ratings. During the study, on 6% of the BRS two separate ratings of the child were made to assess reliability during the study. Agreement within one rank (e.g., 3 vs. 4) occurred on 84% of the items, and within two numbers (e.g., 3 vs. 5) on 99.5% of the items.

2.3.6. Socioeconomic status

Four measures of SES were used: mother's years of school, father's years of school, possessions score, and the floor score. The possessions score was the number of possessions the family owned from a list of 10 common, somewhat significant expenditures. These were car, blender, stove, sewing machine, motorcycle, radio, refrigerator, clock, television, and bicycle. The floor score consisted of the interviewer's observation of the type of floor material on a five-point scale. These were, from lowest to highest, dirt, mixed cement, mosaic, wood, and tile. Most houses had the first two types of floor. Previous work in Guatemala had suggested that the floor material was a good proxy for house quality [41].

2.3.7. Data analysis

Data were analyzed with SPSS version 6.1 for Windows and PC SAS (SAS Institute, Cary, NC). Baseline characteristics were compared between Coffee and Substitute groups using the chi-square statistic for categorical variables and a two-way analysis of variance (treatment by anemia status) for continuous variables. Analysis of variance (ANOVA) was used to compare changes in outcome variables during the intervention, with the main effects being treatment group (Coffee vs. Substitute) and initial anemia status, and including the interaction term between treatment group and initial anemia. Since age of the child was associated with the BSID scores, it was included as a covariate in the ANOVAs. Finally, multiple regression analyses were performed to examine the association of maternal ingestion of coffee during pregnancy with BSID measures controlling for potentially confounding variables.

3. Results

3.1. Sample and characteristics of subjects

The final sample for the assessment of the behavioral effects of coffee was 132 children from the initially recruited 160 subjects. A total of 21 subjects or 13% did not complete the study, 20 because the family moved out of the area, and one due to

hospitalization. No significant differences in socioeconomic status or maternal characteristics were found between those who completed the study and those who did not.

Two subjects had MDI and PDI scores, but did not have data for the BRS, and are included in the MDI and PDI analyses only. Another seven children were lost to the BSID II follow-up: four children who had the final blood draw refused to be tested, and two tests could not be used (one due to an incorrect age, and the other because the child refused to complete the test). One additional subject was dropped from the analyses because she had a MDI score of 50, 11 points below the next lowest score, and a history of epileptic seizures that compromised her ability to be assessed (her PDI score was the third lowest). All of the analyses were run both with and without this child, and her exclusion did not affect the results from the test of the main hypothesis. In the final sample there were 95 children with initial Hgb ≥ 10.5 g/dl and 37 children with an initial Hgb < 10.5 g/dl. Sleep data for the last visit were missing for four subjects, so the sleep variables were available only for 128 children.

Characteristics of subjects who completed the BSID II are shown in Table 1, by initial anemia status and treatment group. The random assignment procedures resulted in generally similar characteristics between intervention groups (within 'non-anemic' or anemic subgroups), except for one of the indices of socioeconomic status (the floor score), which was significantly higher in the Substitute group than in the Coffee group within the 'non-anemic' subgroup. This variable was marginally associated with the PDI, and with the Orientation and Engagement Scale. Therefore, it was used as a covariant in the test of the main effects.

Average age of the children at the beginning of the study was 17 months. The initial *z*-scores indicate a high prevalence of stunting, with an average length-for-age more than 2 SD below the NCHS median. About half of the children were still being

Table 1
Subject characteristics by intervention group and anemia status (mean±SD)

n:	Non-anemic coffee 42	Non-anemic substitute 53	Anemic coffee 18	Anemic substitute 19
Mother's school (years)	4.0 (3.4)	5.0 (3.3)	5.1 (3.4)	5.2 (3.7)
Father's school (years)	6.2 (3.3)	5.5 (3.3)	6.1 (2.8)	6.6 (2.4)
Possessions score	5.0 (1.7)	5.2 (1.7)	4.4 (1.6)	5.0 (2.1)
Floor score	1.6 (0.7)	1.9[a] (0.7)	1.6 (0.7)	1.8 (0.7)
Child age at testing	22.3 (3.8)	23.0 (4.0)	21.2 (3.7)	22.9 (3.4)
Child gender (% F)	48	60	44	47
Birth weight (kg)	3.1 (0.5)	3.1 (0.6)	2.9 (0.5)	2.7 (0.8)
Weight for age *z*-score	− 1.4 (0.9)	− 1.4 (0.9)	− 1.4 (1.1)	− 1.5 (0.8)
Length for age *z*-score	− 2.0 (1.2)	− 2.2 (1.3)	− 2.3 (1.3)	− 2.0 (1.0)
Weight for length *z*-score	− 0.3 (1.1)	0.0 (0.9)	0.0 (1.3)	− 0.4 (1.2)
Initial caffeine/kg/day[b]	8.8 (3.4)	9.9 (17.8)	11.0 (15.2)	6.1 (9.2)
Age began coffee (months)	8.2 (3.4)	7.4 (3.8)	8.3 (3.7)	7.0 (3.2)
Still breast-fed (%)	56	45	37	47

[a]Differs between Coffee and Substitute, $P < 0.05$.
[b]Median and range, respectively, are 20 (1, 120), 21 (3, 227), 28 (0, 126), and 18 (2,76).

breast-fed. The average age at which coffee was introduced was 7–8 months, with 27% of the sample receiving coffee prior to 6 months of age. The majority of families used ground rather than instant coffee. Very few of the children received milk with their coffee. Initial coffee intake averaged 160–200 ml/day and did not differ significantly among groups. Caffeine concentration ranged from 7 to 56 mg/g of dry coffee. Initial caffeine intake averaged 6.1–11.0 mg/kg/day, with a mean of 9.1 mg/kg/day.

3.2. Compliance with the intervention

In general, compliance with the intervention was good. However, in the Substitute group, not all of the children completely discontinued coffee intake: 28% continued to drink coffee (usually in addition to the substitute) during at least three of the ten 2-week intervals during the intervention period (defined as 'non-compliers'). A few others in this group did not accept the substitute consistently, but did not return to drinking coffee. In the Coffee group, there were some children who discontinued coffee intake: 20% reportedly did not drink coffee during at least four of the 2-week intervals (also defined as 'non-compliers'). Mothers had no reason not to answer truthfully to the question about whether they had given coffee, since they were told that they would not be dropped from the study if they were to give the child coffee.

No significant differences between groups in initial food intake or in the change in frequency of consumption of each food type during the intervention were found [35]. There were also no significant differences between Coffee and Substitute groups in initial mean values for Hgb, Hct, ZPP/H, and ferritin [36].

3.3. Effects of the intervention on BSID scores

3.3.1. Measurement issues

Bayley MDI and PDI scores were calculated based on the published age-based standardized norms, because children's ages at the time of testing ranged from 17 to 30 months. The standardized scores were normally distributed according to the Kolmogorov–Smirnov one-sample test of normality. Scores did not differ by gender, nor did they differ according to which of the three testers administered the test. Both scores were negatively associated with child age, even though the score was standardized using age-based normative data, suggesting that the children were falling behind the US norms with age. Therefore, all analyses controlled for the child's chronological age.

Scores for the BRS were based on the sum of the rankings for the items in each factor, and for the total score. Simple summary scores, rather than percentile scores, were used because the former were normally distributed, whereas the percentile score for the two subscales and the total scale were all non-normal. Using a percentile score would not have standardized the scores by age because all of the children were in the same normative group (13–42 months). Positive correlations with age were found. Three scores were used: Orientation and Engagement, Emotional Regulation, and the

Table 2
Partial correlations between BSID scores and SES, anthropometric and hematological indicators controlling for child's age ($N = 132$)

Variable	MDI	PDI	Emotional Regulation score	Orientation Engagement score	Total behavior rating scale
Mother's years of school	0.15[a]	0.04	− 0.01	0.05	0.03
Father's years of school	0.11	0.16[a]	0.18*	0.26**	0.26**
Number of possessions	0.05	0.14	0.05	0.12	0.08
Floor score	0.03	0.16[a]	0.01	0.16	0.14
Parity of child	0.01	0.02	− 0.02	0.03	0.03
Initial length for age *z*-score	0.05	0.15[a]	− 0.06	− 0.12	− 0.11
Initial weight for age *z*-score	0.15[a]	0.27**	− 0.01	0.02	0.02
Initial weight for length *z*-score	0.16[a]	0.17*	0.04	0.09	0.09
Birth weight	0.13	0.23**	0.11	0.11	0.13
Initial ZPP/H[b]	0.00	0.05	− 0.28**	− 0.20*	− 0.25**
Initial Hgb	0.02	0.13	0.36**	0.29**	0.33**
Final ZPP/H[b]	− 0.12	− 0.06	− 0.07	− 0.11	− 0.09
Final Hgb	0.08	0.20*	0.14	0.03	0.08

[a]$P < 0.10$; *$P < 0.05$; **$P < 0.01$.
[b]Values are transformed by taking the natural log. The lower the score, the higher the child's iron status. Thus a negative correlation means that there is a positive association between iron status and BSID II indicator.

Total Score. The Motor Quality score was not used because there was too little variance in the score.

Table 2 presents correlations of socioeconomic status, anthropometric and hematological indicators with the BSID measures controlling for age of the child. Associations with standard SES measures were present but not strong; the MDI was marginally associated with mother's education, PDI scores were marginally related to father's schooling and the floor score, and BRS scales were associated with more education of the father. The MDI and PDI were associated with initial height and weight, and birth weight, but the BRS was unrelated to anthropometric measurements. On the other hand, the BRS was significantly associated with initial indicators of iron status, whereas the MDI and PDI were not.

3.3.2. Treatment effects

The means of the MDI and PDI scores, the BRS scores for Emotional Regulation and Orientation and Engagement, and the Total Score by intervention group and initial anemia status adjusting for child age and floor score are shown in Table 3. There were no significant differences on any of the BSID measures as a function of intervention. MDI and PDI scores did not differ by anemia category; however, the anemic children scored significantly lower on all three BRS ratings. Findings did not differ when the sample was limited to children who complied, or when only children who had initial coffee levels of > 100 ml/day were included. The findings were similar when anemic and non-anemic children were combined, and the resulting two groups were compared by treatment using *t*-tests.

Table 3
Adjusted mean values and analysis of covariance of Bayley Scales of Infant Development II and sleep variables by intervention groups and initial anemia status[a]

n[b]:	Coffee non-anemic 42	Substitute non-anemic 53	Coffee anemic 18	Substitute non-anemic 19	$F_{treatment}$	F_{anemia}	$F_{interaction}$
BSID scores							
MDI	87.6	85.4	89.0	84.6			
PDI	91.6	90.4	91.0	86.6			
Emotional Regulation	38.4	38.6	36.6	34.9		9.58**	
Orientation and Engagement	30.5	29.8	27.5	27.3		8.86**	
Total BRS	102.6	102.4	99.0	96.4		9.66**	
Sleep							
Hours slept total, initial visit	11.3	11.4	11.6	11.3			
Minutes slept naps, initial visit	95.1	84.5	98.8	90.7			
Hours slept night, initial visit	9.7	9.9	10.1	9.7			
Hours slept total, visit 10	10.8	11.3	10.8	11.3	4.16*		
Minutes slept in naps, visit 10	57.8	61.1	64.4	64.4			
Hours slept at night, visit 10	9.9	10.3	9.7	10.3	4.20*		

[a]Controlling for child age and floor score.
[b]Sample sizes for sleep variables in initial visit are slightly smaller.
Note: only F values for significant individual predictors are shown. All models were significant (including covariates).
*$p < 0.05$; **$p < 0.01$.

3.4. Effects of the intervention on sleep variables

Typical sleep patterns at the first visit, prior to the initiation of the coffee or substitute condition, were for children to nap during the day (90%), with 11% taking two naps per day. These naps were primarily in someone's bed (70%) rather than a crib (10%) or a hammock (6%), with night sleeping usually in a bed with someone else. Houses tend to be small, with one or two rooms for the family, and children are put to sleep in a room along with many others. Most parents reported little difficulty getting their children to sleep (79% said very little or no problem), and 70% reported that children fell asleep in 10 min or less. Techniques to get children to sleep were primarily breast feeding (37%), giving the child a bottle and leaving her (22%), giving the child a bottle and sleeping with him (18%), lullabies or rocking (10%) or nothing (8%). Fifty-seven percent of children awoke during the night, and were usually put to sleep with breast feeding or a bottle. Of those who awoke, half awoke once, 26% awoke twice, and 25% awoke three times or more. None of these variables were associated with the child's age or the number of the child's siblings.

At the initial visit, prior to the intervention, no differences by treatment or anemia status were found for the sleep variables: total nap time, total sleep time at night, total sleep in past 24 h, child's difficulty going to sleep, and number of times the child awoke during the night. However, at the final visit, the Substitute children slept significantly more at night and slept more time in total than the Coffee children (Table 4). No differences in total nap time, sleep difficulty or number of night

Table 4
Slopes (*B*) from multiple regression analyses of BSID II variables on maternal coffee intake during pregnancy, with and without controlling for ZPP/H[a] ($N = 129$)

	Maternal coffee intake during pregnancy	ZPP/H[a]	Maternal coffee intake during pregnancy controlling for ZPP/H[a]
MDI	− 0.86	− 0.01	− 0.79
PDI	− 0.41	1.63	− 0.54
Emotional Regulation	− 0.98*	− 2.32**	− 0.30
Orientation and Engagement	− 1.00*	− 1.87*	− 0.54
BRS total	− 0.34[b]	− 3.81**	− 0.68

Note: in all regressions, *B* values shown are after controlling for mother's education, father's education, child's age, and the possessions score. All regression models were significant.
[a]Transformed with a natural logarithm. Since a lower score for ZPP/H indicates higher iron status, negative associations of ZPP/H with BRS mean that higher iron status is associated with higher BRS scores.
[b]$P < 0.10$; *$P < 0.05$; **$P < 0.01$.

wakings were found, and none of the final sleep variables differed by anemia status. The significance of the findings was unchanged when controlling for sleep reported at the initial visit, although the sample size decreased, since 17 children were missing data for the first visit. Differences in sleep time between Coffee and Substitute groups were greatest at the last visit. For the nine earlier visits (time points) during the intervention, marginal differences were found at five time points for total sleep (three favoring the Substitute group) and one time point for total sleep at night (favoring the Substitute group).

3.5. Maternal coffee intake during pregnancy

Women reported ingesting a median of three cups of coffee a day during pregnancy, with a range of none to eight cups. If one assumes that the women weighed an average of 56.4 kg [42] then the average dose of caffeine per day ingested by the women based on the assumptions about cup size and coffee type was almost identical to the average for US women (Guatemalans 2.32 mg/kg/day, US 2.31 mg/kg/day), although the 90th centile and 99th centile levels for Guatemalans were considerably higher than those for the US data [29]. Guatemalan women' reported coffee ingestion during pregnancy was unrelated to any of the SES measures, parity, age of the mother, or age of the child.

In order to assess the association of maternal coffee intake with the Bayley scores, multiple regression for the BSID measures were conducted controlling for the possessions score, maternal and paternal education, and child's age, despite the lack of association of coffee intake with these measures. Maternal coffee ingestion during pregnancy was significantly negatively associated with the Emotional Regulation scale, Orientation and Engagement scale, and Total BRS, but was unrelated to the MDI and PDI (see Table 4).

Dewey et al. [35] found that maternal coffee ingestion during pregnancy was

associated with lower iron status (higher levels of ZPP/H) in the child. Here we examined whether its association with the iron status of the child could explain the association of maternal coffee ingestion during pregnancy with lowered BRS scores. As Table 4 shows, lower log ZPP/H levels in the child (higher iron status) were significantly associated with higher BRS measures. When log ZPP/H was controlled, maternal coffee ingestion during pregnancy was unrelated to BRS scores, suggesting that prenatal coffee intake may influence BRS scores through affecting fetal iron status.

4. Discussion

In this study, the effects of coffee on cognitive development, behavior ratings, and sleep were assessed through a randomized trial, which is a stronger design than most investigations of coffee effects have employed. In observational studies, one has to rely on accurate measurement of coffee intake. Obtaining good data on the dose of coffee/caffeine can be difficult; there are differences in cup size, type (or strength) of coffee, type of beans, and method of preparation that can affect caffeine levels [9].

No significant effects of the treatment on the BSID scores (MDI, PDI or the BRS) were found. There could be several alternative explanations for these negative results. First, the BSID II measure might not have been appropriate for the study population. However, although the BSID II had not previously been standardized in Guatemala, it appeared to be appropriately associated with criteria; the scores correlated with socio-economic status variables similarly to other studies [43]. The PDI and MDI were correlated with anthropometric status, a pattern often reported [43,44]. Also, the mean score for the Mental Development Index of 86.6 (SD = 10.0) was very similar to that found in a group of inner city Black children tested in Baltimore [45].

A second possible reason for the absence of effects is that the amounts of coffee ingested may have been too low to have an effect on development. However, when translated into caffeine intake per kilogram of body weight, the average consumption of caffeine by Guatemalan infants (9 mg/kg/day) is very close to the top 1% of caffeine consumption reported in the US, over three times the amount that adults in the US ingest per kilogram, and over six times the average ingestion of American 2–17 year olds [29]. Sobotka [29] found that US infants (6–24 months) ingested < 1 mg of caffeine per kilogram per day, and Arbeit et al. [46] reported 1.2–6.4 mg/kg/day ingestion among children aged 6 months to 3 years in Louisiana. The highest decile of infants (aged 0–1 years) and preschoolers (aged 1–5) in a US sample ingested 0.7 and 2.1 mg/kg/day, respectively [47]. Amounts of caffeine ingested by children appear to be higher in Guatemala. Thus the caffeine intake in the present study was probably sufficient to observe an effect had one occurred.

A third possibility is that the groups were not initially similar in cognitive indices, and without baseline data we cannot eliminate this possibility. However, it seems unlikely given the similarity between the groups on the other characteristics and the use of random assignment. Fourth, the study may have had insufficient power to detect a difference, but the power estimates suggested that it should have been adequate. Thus we are left with the conclusion that 5 months without coffee does not

affect cognitive function in this sample. We had postulated both a direct effect of coffee on behavior, and an indirect effect through changes in iron status, but neither was found. However, as we report elsewhere, coffee restriction did not affect iron status in these children [35]. Given the absence of an effect of coffee restriction on iron status, these insignificant results are less surprising; removing coffee for 5 months did not affect either iron status or Bayley test scores.

Five months may have been insufficient time to see an effect. In a separate report, differences between treatment groups were seen in growth and morbidity [36]. When the sample was limited to children who had initially ingested more than 100 ml/day of coffee, a significant effect of treatment on weight gain was found among the anemic subjects, and a marginal improvement was found in length in the combined anemic and 'non-anemic' subgroups. It is possible that improvements in anthropometric status following discontinuation of coffee might eventually have a long-term effect on cognitive functioning.

Significant differences were found in sleep patterns by treatment group. At the last visit, children who discontinued coffee were found to sleep significantly more during the night, and overall, than children who continued to receive coffee. This finding is consistent with a number of studies in adults that have found that coffee inhibits sleep, even some time after ingesting the coffee [18,30–33]. No differences were found prior to the introduction of the intervention. However, these differences were significant only at the end of the study; marginal differences, primarily in favor of the substitute group, were found in over half of the visits earlier in the study. It is possible that the effects increased over time, but these findings should be investigated further.

In contrast to the lack of effect of postnatal coffee restriction on the BSID II, the mother's coffee intake during pregnancy was inversely associated with the child's behavior during the testing situation, as measured by the Behavior Rating Scale. This association held even controlling for a number of potentially confounding variables. Maternal coffee ingestion was also negatively associated with the child's initial iron status [35]. Similarly, in Costa Rica, coffee intake of mothers during pregnancy was negatively associated with hematological status of the infant at 1 week and 1 month of age [5]. In the present study, multiple regression suggested that the association of maternal coffee ingestion with increased emotionality in the testing situation was due to changes in the child's iron status.

The MDI and PDI are reported more often than the Infant Behavior Record, the rating scale in the first BSID, which measures temperament [48] or behavioral style [49]. The MDI and PDI may be less accurate predictors of later cognitive ability than the behavior ratings. For example, scores from the Infant Behavior Record were related to language skills at ages 3 and 7 years [48], and were better predictors of later IQ scores than the MDI and PDI measures [40,50–54]. Thus the lower BRS scores for the children of mothers ingesting coffee might have consequences for their later cognitive development. Strupp and Levitsky [55] have hypothesized that increased emotional reactivity, reflected in anxiety and upset in new situations, for an extended period of time may have significant consequences for mental development. These investigators suggest that heightened arousal and response to stress may be the most marked response to malnutrition, and can play a major role in a child's

development. Lozoff [56] has proposed a similar link between anemia, emotionality, and cognitive development.

Most of the research involving the effects of coffee or caffeine on cognitive development has been conducted on pregnant mothers and their offspring. Data from both human and animal studies suggests subtle behavioral effects of prenatal coffee intake on offspring, particularly increased reactivity and locomotor activity [57,58]. Because caffeine readily crosses both the blood–brain and the placental barriers [59] it is a potential hazard to the developing fetus. However, frank effects are seen only at higher levels of intake. In their recent review, Nehlig and Debry [57] found that long-term disturbances in animal behavior are evident only after maternal caffeine exposures of 75–120 mg/kg/day, which would be equivalent to more than 50 cups of coffee a day in humans. On the other hand, Purves and Sullivan [60] suggested that prenatal exposures of more than 20 mg/kg/day of caffeine could cause behavioral alterations in animals, although the effects were subtle and somewhat inconsistent.

Studies of the effect of human mothers' caffeine consumption on infants' behavior have had mixed results. Streissguth and colleagues found evidence for a long-term effect of alcohol, but not nicotine or caffeine, in children assessed at 8 months [61], 4 years of age [62], and 7 years [63]. Maternal consumption of 4.8 mg of caffeine/kg body weight/day during pregnancy resulted in moderate plasma concentrations of caffeine, but did not affect newborns' Apgar scores [59]. Jacobson et al. [64] found maternal caffeine exposure of 0.5–8.5 mg/kg/day adversely effected neonatal behavior, reflex functioning, and neuromuscular development on the Brazelton Neonatal Behavior Assessment Scale. These effects appeared to be independent of confounds such as nicotine, alcohol, demographic variables, stress, and obstetrical medication. Hinds et al. [65] conclude from a literature review of primarily US data that heavy caffeine use ($\geq$300 mg per day, or about 4 mg/kg/day for a woman of 160 lbs.) during pregnancy is associated with small reductions in infant birth weight. Therefore, this area should receive more attention. In the US population heavy coffee drinkers may be more likely to smoke which makes interpretation of data more complex. However, in Guatemala very few women smoke, allowing one to separate the effects of coffee from those due to smoking or other life-style differences.

In this study maternal coffee consumption during pregnancy was negatively associated with the child's iron status, and with the child's emotionality and responsiveness during testing. These data concur with Hind et al.'s [65] conclusion that prenatal intake of coffee should continue to be a concern, and deserves the same experimental attention afforded to the postnatal effects. Based on effects observed in rodents, in 1980 the US Food and Drug Administration published a warning recommending that pregnant women limit or even eliminate coffee consumption [66]. Additional experimental studies are needed to evaluate possible causal relationships between maternal coffee intake during pregnancy and child outcomes.

Acknowledgements

This work was supported by a grant from the Thrasher Research Fund. We appreciate the mothers and children who participated in the study, and the field

support of the psychological testers, Mirna de la Corona Vela, Liliana Aldana, and Ester Rodriguez, Cal Poly student interns Yvonne Corral-Koch and Ivan Vela, and the field team of Rina Pena, Guadalupe Juárez, Karin Casasola, Carolina González, Lucrecia Aldana, Edna López, Virginia Mazariegos and Azucena Méndez.

References

[1] The Food and Agriculture Organization. Food balance sheets. Rome: The Food and Agriculture Organization, 1989.
[2] Engle PL, Pedersen ME. Maternal work for earnings and children's nutritional status in rural urban Guatemala. Ecol Food Nutr 1989;22:211–23.
[3] Valverde B, Arroyave G. Revision de aporte calorico y proteinico de las dietas de poblaciones de bajo nivel socioeconomico en Centroamerica. Arch Latinoamericanos Nutricion 1975;25:327–49.
[4] Quan de Serrano J, Gamero H, Bulux J, Zepeda E, Guerro AJ, Lopez CY, Vasquez A, Vettarazzi L, Dewey K. Introduction of solid and liquid foods in the weaning process of contemporary Guatemalan infants (abstract). FASEB J 1991;4:A1289.
[5] Munoz LM, Lonnerdal B, Keen CL, Dewey KG. Coffee consumption as a factor in iron deficiency anemia among pregnant women and their infants in Costa Rica. Am J Clin Nutr, 1988 Sep;48(3):645–51.
[6] Nakamoto T, Roy G, Gottschalk S, Yazdani M, Rossowska M. Lasting effects of early chronic caffeine feeding on rats' behavior and brain in later life. Physiol Behav 1990;49:721–7.
[7] Marangos PJ, Boulenger JP, Patle J. Effects of chronic caffeine on brain adenosine receptors: anatomical and ontogenic studies. Life Sci 1984;34:899–907.
[8] Peruzzi G, Lombardelli G, Abbrachio MP, Coen E, Cattabeni F. Perinatal caffeine treatment: behavioral and biochemical effects in rats before weaning. Neurobehav Toxicol Teratol 1985;7:453–60.
[9] Nehlig A, Debry G. Potential teratogenic and neurodevelopmental consequences of coffee and caffeine exposure: a review on human and animal data. Neurotoxicol Teratol 1994;16:531–43.
[10] Rapoport J, Berg C, Ismond D, Zajn T, Neims A. Behavioral effects of caffeine in children. Arch Gen Psychiatry 1984;41:1073.
[11] Baer RA. Effects of caffeine on classroom behavior, sustained attention, and a memory task in preschool children. J Appl Behav Anal 1987;20:225–34.
[12] Rapaport JL, Jensvold M, Elkins R. Behavioral and cognitive effects of caffeine in boys and adult males. J Nerv Mental Dis 1981;169:726–32.
[13] Bernstein GA, Garfunkel BD. Caffeine effects on learning, performance, and anxiety in normal school-aged children. J Am Acad Pediatr 1994;33:407–15.
[14] Fagan D, Swift CG, Tiplady B. Effects of caffeine on vigilance and other performance tests in normal subjects. J Psychopharmacol 1988;2:19–25.
[15] Anderson KJ, Revelle W, Lynch MJ. Caffeine, impulsivity, and memory scanning: a comparison of two explanations for the Yerkes–Dodson effect. Motivation Emotion 1989;13:1–20.
[16] Gupta U. Effects of impulsivity and caffeine on human cognitive performance. Pharmacopsychoecologia 1988;1:33–41.
[17] Dulloo AG, Geissler GA, Horton T, Collins A, Miller DS. Normal caffeine consumption: influences on thermogenic and daily energy expenditure in lean and post-obese human volunteers. Am J Clin Nutr 1989;49:44–50.
[18] Bonnet MH, Arand DL. Caffeine use as a model of acute and chronic insomnia. Sleep 1992;15(6):526–36.
[19] Hallberg L, Rossander L. Effect of different drinks on the absorption of non-heme iron from composite meals. Hum Nutr: Appl Nutr 1982;36:116–23.
[20] Morck TA, Lynch SR, Cook JD. Inhibition of food iron absorption by coffee. Am J Clin Nutr 1983;37:416–20.

[21] Oski FA, Honig AS. The effects of therapy on the developmental scores of iron deficient infants. J Pediatr 1978;92:21–5.
[22] Lozoff B, Brittenham GM, Viteri FW, Wolf AW, Urrutia JJ. The effects of short-term oral iron therapy on developmental deficits in iron-deficient anemic infants. J Pediatr 1982;100:351–7.
[23] Lozoff B, Brittenhanm GM, Wolf AW. Iron deficiency anemia and iron therapy effects on infant developmental test performance. Pediatrics 1987;79:981–95.
[24] Aukett MA, Parks YA, Scott PH, Wharton BA. Treatment with iron increases weight gain and psychomotor development. Arch Dis Child 1986;61:849–57.
[25] Idjradinata P, Pollitt E. Reversal of developmental delays in iron-deficient anemic infants treated with iron. Lancet 1993;341:1–4.
[26] Pollitt E, Kim I. Learning and achievement among iron-deficient children. In: Youdim MBH, editor. Brain iron: neurochemical and behavioral aspects. Topics in Neurochemistry and Neuropharmacology. Taylor and Francis, 1988.
[27] Heywood A, Openheimer S, Heywood P, Jolley D. Behavioral effects of iron supplementation in infants in Mandang, Papua New Guinea. Am J Clin Nutr 1989;50(Suppl):630–40.
[28] Lozoff B, Jimenez E, Wolf A. Long-term developmental outcome of infants with iron deficiency. N Engl J Med 1991;325:687–94.
[29] Sobotka TJ. Neurobehavioral effects of prenatal caffeine. Ann NY Acad Sci 1989;562:327–39.
[30] Nicholson AN, Stone BM. Heterocyclic amphetamine derivatives and caffeine on sleep in man. Br J Clin Pharmacol 1980;9:195–203.
[31] Karacan I, et al. Dose-related sleep disturbances induced by coffee and caffeine. Clin Pharmacol Ther 1976;20:682–9.
[32] Nehlig A, Daval J, Debry G. Caffeine and the central nervous system: mechanisms of action, biochemical, metabolic and psychostimulant effects. Brain Res Rev 1992;17(2):139–69.
[33] Walsh JK, Muelhbach MJ, Humm TM, Dickins QS, Sugerman JL, Schweitzer PK. Effect of caffeine on physiological sleep tendency and ability to sustain wakefulness at night. Psychopharmacology 1990;101:271–3.
[34] Aranda JV, Collinge JM, Zimman R, Watters J. Maturation of caffeine elimination in infancy. Arch Dis Child 1979;54:946–9.
[35] Dewey KG, Romero-Abal E, Quan de Serrano J, Bulux J, Peerson JM, Engle PL, Solomons N. Effects of discontinuing coffee intake on iron status of Guatemalan toddlers: a randomized intervention study. Am J Clin Nutr 1997;66:168–76.
[36] Dewey KG, Romero-Abal ME, Quan de Serrano J, Bulux J, Peerson JM, Engle PL, Solomons N. Effects of discontinuing coffee intake on growth and morbidity of Guatemalan toddlers: a randomized intervention study. J Nutr 1997;127:306–13.
[37] Dallman PR. Biochemical and hematological indices of iron deficiency. In: Pollitt E, Leibel RL, editors. Iron deficiency: brain biochemistry and behavior. New York: Raven Press, 1982:63–78.
[38] The Psychological Corporation. The Bayley Scales of Infant Development II. San Antonio, TX: Harcourt, Brace and Company, 1993.
[39] Wolf AW, Lozoff B. A clinically interpretable method for analyzing the Bayley Infant Behavior Record. J Pediatr Psychol 1985;10(2):199–214.
[40] Matheny AP, Wilson RW. Developmental tasks and rating scales for the laboratory assessment of infant temperament (ms. No. 2367). JSAS Catalog of Selected Documents in Psychology 1981;11:81–2.
[41] Engle PL, Carmichael SL, Gorman K, Pollitt E. Demographic and socio-economic changes in the Guatemalan Oriente families from 1967 to 1987. Food Nutr Bull 1993;14(3):237–45.
[42] Villar J, Dorgan J, Menendez R, Bolanos L, Pareja G, Kestler E. Perinatal data reliability in a large teaching obstetric unit. Br J Obstet Gynaecol 1988;95:841–8.
[43] Johnston F. The urban disadvantage in the developing world and the physical and mental growth of children. In: Scell LM, Smith MT, Bilsborough A, editors. Urban ecology and health in the Third World. Cambridge, UK: Cambridge University Press, 1993:26–37.
[44] Lasky RE, Klein RE, Yarbrough C, Engle PL, Lechtig AL, Martorell LR. The relationship between physical growth and infant behavioral development in rural Guatemala. Child Dev 1981;52:219–26.
[45] M.M. Black. Personal communication, 1996.

[46] Arbeit ML, Nicklas TA, Frank GC, Webber LS, Miner MH, Berenson GS. Caffeine intakes of children from a biracial population: the Bogalusa Heart Study. J Am Diet Assoc 1988;88(4):466–71.
[47] Barone JJ, Roberts HR. Caffeine consumption. Paper presented at the American Academy of Pediatrics Meeting, October 27, 1991, New Orleans, LA.
[48] Slomknowski CL, Nelson K, Dunn J, Plomin R. Temperament and language: relations from toddlerhood to middle childhood. Dev Psychol 1992;28(6):1090–5.
[49] Baroni MA. Bayley's Infant Behavior Record ratings of infants with recurrent apnea: behavioral profile and correlates with apnea, age, and developmental status. J Dev Behav Pediatr 1992;13(3):158–64.
[50] Matheny AP. Bayley's Infant Behavior Record: behavioral components and twin analyses. Child Dev 1980;51:1157–67.
[51] Matheny AP. A longitudinal twin study of stability of components from Bayley's Infant Behavior Record. Child Dev 1983;54:356–60.
[52] Kaplin-Estrin M, Jacobson SW, Jacobson JL. Alternative approaches to clustering and scoring the Bayley Infant Behavior Record. Infant Behav Dev 1994;17(2):149–57.
[53] Benson B, Cherny SS, Haith MM, Fulker DW. Rapid assessment of infant predictors of adult IQ: midtwin–midparent analyses. Dev Psychol 1993;29(3):434–47.
[54] Cardon LR, Fulker DW. Sources of continuity in infant predictors of later IQ. Intelligence 1991;15(3):279–93.
[55] Strupp BJ, Levitsky DA. Enduring cognitive effects of early malnutrition: a theoretical reappraisal. J Nutr 1995;125:2221S–32S.
[56] Lozoff B. Explanatory mechanisms for poorer development in iron-deficient anemic infants. In: Nutrition, health, and child development: research advances and policy recommendations. Pan American Health Organization. Tropical Metabolism Research Unit of the University of the West Indies, and The World Bank. Scientific Publication 566, 1998:162–178.
[57] Nehlig A, Debry G. Consequences on the newborn of chronic maternal consumption of coffee during gestation and lactation: a review. J Am Coll Nutr 1994;13(1):6–21.
[58] Leviton A. Coffee, caffeine, and reproductive hazards in humans. In: Garattini S, editor. Caffeine, coffee and health. New York: Raven Press, 1993:343–358.
[59] Van't Hoff W. Caffeine in pregnancy. Lancet 1982;1:1020.
[60] Purves D, Sullivan FM. Reproductive effects of caffeine. Experimental studies in animals. In: Garattini S, editor. Caffeine, coffee, and health. New York: Raven Press, 1993:317–342.
[61] Streissguth AP, Barr HM, Martin DC, Herman CS. Effects of maternal alcohol, nicotine, and caffeine use during pregnancy on the infant mental and motor development at eight months. Alcohol Clin Exp Res 1980;4:152–64.
[62] Streissguth AP, Barr HM, Martin DC, Herman CS. Intrauterine alcohol and nicotine exposure: attention and reaction time in 4-year-old children. Dev Psychol 1984;20:533–44.
[63] Barr HM, Streissguth AP. Caffeine use during pregnancy and child outcome: a 7-year prospective study. Neurotoxicol Teratol 1991;13:441–8.
[64] Jacobson SW, Fein GG, Jacobson JL, Schwartz PM, Dowler JK. Neonatal correlates of prenatal exposure to smoking, caffeine, and alcohol. Infant Behav Dev 1984;7:253–65.
[65] Hinds TS, West WL, Knight EM, Harland BF. The effect of caffeine on pregnancy outcome variables. Nutr Rev 1996;54(7):203–7.
[66] Berger A. Effects of caffeine consumption on pregnancy outcome. J Reprod Med 1988;33:954–6.