Person. individ. Diff. Vol. 12, No. 4, pp. 351-362, 1991
Printed in Great Britain

0191-8869/91 $3.00 + 0.00
Pergamon Press plc

CONTROLLED TRIAL OF VITAMIN-MINERAL SUPPLEMENTATION: EFFECTS ON INTELLIGENCE AND PERFORMANCE*

S. J. SCHOENTHALER,[1] S. P. AMOS,[1] H. J. EYSENCK,[2] E. PERITZ[3] and J. YUDKIN[4]

[1]Department of Sociology and Criminal Justice, California State University, Stanislaus, Turlock, CA 95380, U.S.A., [2]Institute of Psychiatry, University of London, London SE5 8AF, England, [3]University of Jerusalem, Jerusalem, Israel and [4]Kings College London, Campden Hill Road, London W8 7AH, England

Summary—615 schoolchildren were examined on multiple measures of intelligence, randomly assigned to one of four treatment groups for 12 weeks, and post-tested on the same measures of intelligence. One cohort received placebos while the other three were given different strength vitamin-mineral supplements. The trial was completed under "blind" conditions, i.e. the subjects, the testers, and the scientists conducting the data analyses did not know any subject's group assignment. This created a triple-blind placebo-controlled, classical design capable of determining whether supplements could produce significant gains in standardized validated indices of intelligence/performance. The study was carried out in Stanislaus County, California, using 4 different schools. Results showed that for non-verbal Wechsler Tests there were highly significant improvements in I.Q., whereas for the verbal tests there were none; a conclusion expected on the basis that "fluid" intelligence, as measured by the non-verbal test, might be improved by supplementation, wheras "crystallized" ability tests (verbal tests) would be unlikely to be so improved. This difference was predicted on the basis of previous studies also finding similar results. Other tests (Raven's Matrices, the Comprehensive Test of Basic Skills, the Matrix Analogies Test, and measures of Reaction Time and Inspection Time) gave additional confirmatory evidence on the contribution which supplementation of the diet by vitamins and minerals can make to the improvement of I.Q.

INTRODUCTION

In 1986, a 16% gain in academic performance was reported for several hundred New York schools following implementation of diet policies that sharply lowered sugar, fat and additive consumption (Schoenthaler, Doraz & Wakefield, 1986a, b). The academic gains and diet changes did not appear to be independent; just before the diet changes, the children showed a significant decline in academic performance in the 133 schools where the most cafeteria meals were served. However, after the diet changes the same schools not only produced gains, but produced the largest gains in academic performance among the 803 schools in the district ($P < 0.001$). The authors hypothesized that if the diet changes were connected with the changes in performance, then the primary cause was probably the uptake of vitamins and minerals in the more nutritious metals rather than an unknown toxic property of sugar, fats and additives.

In 1988, the British Broadcasting Corporation televised a documentary on two independent controlled trials that examined the vitamin-mineral supplementation and intelligence hypothesis. Both studies reported a significantly larger gain in non-verbal I.Q. among children on supplements than children on placebos. Neither found a change in verbal I.Q. The results seemed to suggest that improving nutrition can raise non-verbal intelligence.

This difference between improvement on verbal and non-verbal tests is of vital theoretical importance. Verbal tests are largely measures of "crystallized" ability, non-verbal tests of "fluid" ability, as defined by Cattell, Horn and others (Eysenck, 1979; Matarazzo, 1972). We would not expect nutritional supplements to affect crystallized ability, as measured, say, by vocabulary tests; if there are any effects, we would expect them to be manifested in tests of fluid ability. This difference, in fact, would constitute an important part of the theory tested.

Sharp debate and inadequate replications followed their release. The Dietary Research Foundation (DRF), a non-profit licensed British charity, was created in 1989 to fund and oversee two well-controlled independent trials designed to test two hypotheses: First, will schoolchildren placed

*A smaller replication of this study was carried out in Great Britain, under the supervision of Dr D. Tamir of Jerusalem. Results were generally supportive of those reported here and will be published in due course. The Iraqi War caused the call-up of Dr Tamir, who was thus unable to complete the account of the British study.

on supplements produce a significantly greater increase in non-verbal I.Q. than children placed on placebos? Second, will the proportion of children on supplements whose gains in non-verbal I.Q. exceeded 14 points be significantly greater than the proportion of children on placebos who produce similar gains? This paper discusses the results of the larger DRF study that examined these hypotheses with American schoolchildren.

LITERATURE REVIEW

The theoretical support

Orthodox malnutrition theory. Most nutritionists and physicians would not expect that giving supplements in western countries could involve brain function, i.e. intelligence. In order for malnutrition to impair brain function, the brain's nutrient source, the blood stream, would have to first experience a pathological decline in nutrient concentrations. Prolonged low intake and tissue depletion manifest themselves as low blood concentrations and physical (i.e. clinical) signs of malnutrition (Underwood, 1986). However, low blood nutrient concentrations are not believed to be widespread in western nations since physical signs of malnutrition are not common.

An alternative theory. A closer review of the literature shows that brain function is affected *before* physical signs and symptoms appear for selected nutrients (Sandstead, 1986; Cherkin, 1987). Thus, one should not assume that subclinical malnutrition is not present because of the absence of physical signs and symptoms. Furthermore, it is important to note that current definitions of adequate nutrition are based on physical conditions rather than mental function. What is adequate to prevent physical signs and symptoms may not be adequate to prevent impaired mental function.

The authors expected that a supplement would have no more effect than a placebo on the intelligence of children who do not have nutritional deficiencies. However, there must exist an unknown number of children who are subclinically malnourished. It is this group of unknown size that is hypothesized to respond to supplementation.

An analogy may be useful. Some children who were earning poor marks improve dramatically after receiving their first pair of glasses. If reading glasses were widely distributed where no one had glasses, a few children would perform much better. Yet, the average increase in performance would be quite small since most children did not need glasses. But for those children with sight problems, the changes should be dramatic.

Likewise one could hypothesize that, if supplements were widely distributed, the typical student would gain no more than 2 to 5 points in I.Q. since most children should be sufficiently well nourished. However, a sub-sample should produce gains that match those found in children who needed glasses. For them, gains of 10 or more points in I.Q. should not be unusual. It follows that a fair test of the link between non-verbal intelligence and nutrition must either: (1) compare mean gains in I.Q. among children on placebos and supplements with a sample size that is sufficient to detect a real difference of only 2 points; or (2) compare the proportion of children on placebos and supplements who produce "dramatic" gains. The second hypothesis requires a much smaller sample. Both hypotheses have been tested in this study.

The empirical support

Early studies. It is well known that the mental performance of children who were severely malnourished during infancy is significantly worse than well-nourished children (Cravioto & Delicardie, 1970). Reaction time and visual–perceptual assignments such as simple copying are impaired in these children. Unfortunately, it is difficult to separate how much of the difference is due to malnutrition or low social economic status since the two are usually found together.

Cravioto randomly assigned pregnant women to one of three groups receiving different strengths of supplements. The study showed that the infants from the group that received the greatest variety of essential nutrients produced significantly better intellectual functioning 12 months after birth. More specifically, the speed at which the babies became used to a figure on a screen, as indicated by looking away, and the speed which they become interested again when a new figure appeared was significantly better among the group that received higher dose vitamins, minerals, and protein (Rush, Stein & Susser, 1980). A review of the studies that show a link between supplements and intelligence, skipping breakfast and I.Q., sucrose and intelligence, food additives and intelligence, toxic metals and intelligence, and general malnutrition and behavior has been reported elsewhere and is beyond the scope of this paper (Schoenthaler, 1991; Conners, 1989).

Recent studies. Schoenthaler, in the third article in this Symposium, randomly assigned 26 children confined in an American juvenile correctional facility to a placebo or supplement group for 3 months. Each subject had been given the Wechsler Intelligence Scale for Children—Revised (WISC-R), when first confined over a year before. The supplement group produced a 5 point gain in non-verbal I.Q., while the placebo group fell 1 point ($P < 0.05$), but the mean differences are somewhat misleading.

Only 5 of 15 children on supplements produced gains of 9 or more points in non-verbal I.Q., i.e. a change that would exceed the 99% confidence interval. These 5 produced an average gain of 20 points. The other 10 children on supplements showed an average gain of zero points. In marked contrast, only 1 of 11 children on placebos produced a gain in excess of 9 points with the remaining 10 showing an average decline of 1 point. It is noteworthy that the child who received placebos and produced a major gain in I.Q. was the only child in the placebo group that had improved his diet dramatically while confined. Thus, there was no impact on most of the children, with 6 dramatic exceptions. The data seems to suggest that the intellectual response of a few children will be dramatic following improvement of their nutrition with supplements or diets.

Benton and Roberts (1988) assigned 30 British subjects to placebos and 30 to supplements for 10 months. They reported a significant gain in non-verbal intelligence of 10 points with no change in the placebo group. Unlike the American study, the paper did not report the number of children on supplements who produced dramatic gains in contrast to the placebo group. The absence of such data implied that a mean gain of 10 points was typical.

Crombie, Todman, McNeill, Florey, Menzies and Kennedy (1990) attempted to replicate Benton's work with 86 children who were given supplements or placebos for 7 months. Although the active group gained 2.4 points more than the placebo group, the difference was not significant. Crombie pointed out that his study had a power of 84% for detecting a difference of 6 points or more which contributed to his conclusion that no link existed between diet and I.Q. Unfortunately, Crombie *et al.* (1990), Benton and Buts (1990) and Nelson, Naismith, Burley and Gatenby (1988) did not examine the primary hypothesis offered here; only a subset of the population should respond, i.e. the unknown proportion that is malnourished as suggested by Yudkin (1988) in the *Lancet*. For example, Crombie's 42 subjects on supplements (who generated a mean gain of 2.4 units more than the placebo group) must have produced a total difference of 101 units, i.e. 42×2.4. One cannot help but wonder if these 101 units of gain occurred among 10 or less of the 42 students on supplements with few or no corresponding gains in the placebo group? If so, the results would be identical to the reported American findings. Unfortunately, Crombie's sample of 86 children was too small to determine if the difference of 2.4 points between his groups was statistically significant.

Nelson *et al.* (1988) attempt at replication had a superior sample size, i.e. 157 students and might have been able to detect a significant difference. Unfortunately, the length of the trial was very short, 1 month instead of 10. This raised the serious question of how long one must wait before post-testing subjects.

Last, but not least, the formulas used by Benton, Crombie and Nelson were suspect. The number of nutrients and strengths are lower than those used in the American trial. This raised the question of what nutrients and strengths were needed when testing a nutrition and intelligence hypothesis. The DRF resolved these issues by requiring three different supplement formulas in both Great Britain and America and using over 1000 subjects.

RESEARCH METHODS

Classical research design

In America 615 schoolchildren were examined on multiple measures of intelligence, randomly assigned to 1 of 4 treatment groups for 13 weeks, and post-tested on the same measures of intelligence. One cohort received placebos while the other three were given different strength vitamin-mineral supplements. The subjects, the testers, and the scientists conducting the data analysis did not know any subject's group assignment until after the data analyzed.

The measures of intelligence/performance that were given as a pre- and post-test were: (1) the Wechsler Intelligence Scale for Children—Revised (WISC-R); (2) the Matrix Analogies Test (MAT) (Naglieri, 1985); (3) Reaction Time and Inspection Time (RT/IT); and (4) the Comprehensive Test of Basic Skills (CTBS). The Raven Matrices (RM) was also given to see if significant differences appeared after one month of supplementation; this was a single test not a retest.

A number of control and secondary variables were gathered. They included: (1) pre- and post-test blood samples that were assayed for vitamin and mineral content; (2) the degree of compliance with taking the pills; (3) the Eysenck Junior Personality Questionnaire (EPQ); and (4) a number of demographic variables, i.e. age, race, sex, classroom and school.

Informed consents

Parents received letters that explained the nature of the project and requested permission to ask their child to participate. Students with parental consent were given the choice of: (1) not participating; (2) taking pills and selected group tests; or (3) total participation which included the WISC-R and donating a pre- and post-blood sample.

Locations

The study was limited to 8th graders (children aged 12–13 years) and 10th graders (aged 15–16 years) from 4 schools near California State University Stanislaus. The schools represent a cross-section of American social economic classes. Mae Hensley Middle School and Ceres High School are middle class schools in typical middle class neighborhoods. Riverbank High School lies in an economically depressed area. Half the residents are hispanics and half are caucasian with a substantial proportion on public assistance. Oakdale High School's students come from homes that are among the most expensive in the county. Nearly half of Oakdale's students have I.Q.s above 120. They represent the brightest 5% of the population. Thus, the schools represent a broad socio-economic cross-section of America. Participation ranged from 45 to 71% in each school with a mean of 59% as illustrated in Table 1.

Formulas

According to Professor Yudkin it was unlikely that many children were consuming <50% of the USRDAs (i.e. the United States Recommended Daily Allowances). If so they would be showing widespread signs of physical deficiencies. Thus he reasoned that a tablet containing 50% of the USRDA for all the vitamins and minerals should ensure that all participating children were consuming at least 100% of the USRDA each day. It follows that a decision was made to test one formula containing 50% of the USRDA.

On the other hand, the RDAs were not established based on brain function. Bringing every child's daily intake up to 100% of the RDA, by giving a supplement that contained 50% of the USRDA, might be inadequate to eliminate any deficiencies that impaired brain function. Furthermore, Benton and Schoenthaler used doses closer to 100% of the USRDAs. If no relationship was found between I.Q. and pills containing 50% of the RDAs, the strength of the dose could be questioned. Thus, a second formula that contained 100% of the USRDAs for adults became essential. Many scientists have attacked the RDAs of both the U.S. and Great Britain as unreasonably low. Therefore, the decision was made to test an even higher strength supplement set at 200% of the USRDAs for most, but not all vitamins and minerals. The composition of all three formulas is shown in Table 2. The test formulas never exceeded 200% of the RDAs for adults when the pills were made in 1989. The formulas were indistinguishable in size, color, odor, and taste.

Table 1. Sample selection

	Potential subjects	Actual subjects	% Participation
Ceres High School (10th graders)	322	146	45
Oakdale High School (10th graders)	205	102	50
Riverbank High School (10th graders)	100	66	66
Mae Hensley Junior High (8th graders)	424	301	71
Total	1051	615	59

Table 2. Formulas

	Formula One 50% RDA		Formula Two 100% RDA		Formula Three 200% RDA	
Vitamin A	2500 I.U.	50%	5000 I.U.	100%	5000 I.U.	100%
Vitamin D	200 I.U.	50%	400 I.U.	100%	400 I.U.	100%
Vitamin E	15 I.U.	50%	30 I.U.	100%	60 I.U.	200%
Vitamin C	30 mg	50%	60 mg	100%	120 mg	200%
Thiamin	0.75 mg	50%	1.5 mg	100%	3 mg	200%
Riboflavin	0.85 mg	50%	1.7 mg	100%	3.4 mg	200%
Niacin	10 mg	50%	20 mg	100%	40 mg	200%
Pantothenic acid	5 mg	50%	10 mg	100%	20 mg	200%
Pyridoxine	1 mg	50%	2 mg	100%	4 mg	200%
Folic acid	200 mcg	50%	400 mcg	100%	400 mcg	100%
Vitamin B12	3 mcg	50%	6 mcg	100%	12 mcg	200%
Biotin	150 mcg	50%	300 mcg	100%	300 mcg	100%
Vitamin K	50 mcg	*	50 mcg	*	50 mcg	*
Calcium	200 mg	20%	200 mg	20%	200 mg	20%
Magnesium	80 mg	20%	80 mg	20%	80 mg	20%
Iron	9 mg	50%	18 mg	100%	36 mg	200%
Zinc	7.5 mg	50%	15 mg	100%	30 mg	200%
Iodine	75 mcg	50%	150 mcg	100%	150 mcg	100%
Copper	1 mg	50%	2 mg	100%	3 mg	150%
Manganese	1.25 mg	*	2.5 mg	*	5.0 mg	*
Chromium	0.05 mg	*	0.10 mg	*	0.20 mg	*
Selenium	0.05 mg	*	0.10 mg	*	0.20 mg	*
Molybdenum	0.12 mg	*	0.25 mg	*	0.50 mg	*

*The USRDA has not been established for these essential nutrients. However, none of these dose levels exceed the National Academy of Science's "Estimated Safe and Adequate Daily Dietary Intakes of Additional Selected Nutrients".

Pre-tests

Both the verbal and non-verbal portion of the WISC-R were given to the 440 children who agreed to donate blood samples. All 615 volunteers were asked to complete the MAT and EPQ. Blood samples were drawn shortly after the psychological testing. A few students who agreed to donate blood changed their minds or were absent when blood was drawn, causing the number of children who took the WISC-R to exceed the number who donated blood.

Distribution of pills

Supplements and placebos were distributed following completion of the pre-tests for 13 weeks. Each student was supposed to swallow one vitamin pill and one mineral pill Tuesday through Thursday with a double dose on Monday and Friday to make up for the weekend. Whenever children were absent 1 day, they were given a double dose the next day. If they were absent more than 1 day, they were only given 1 double dose the first day back to school. No effort was made to make up for the other missed pills. In theory, each participant took 14 pills each week, but due to absences of more than one day, the participants who completed the study averaged 13 pills per week.

The Raven's Matrices were given after the students had been receiving pills for 4 weeks, i.e. they received only about one-third of the supplementation they received for the other tests.

Post-tests

Project staff gave the WISC-R, the EPQ, and the MAT a second time 10–13 weeks later. Of 411 subjects who took a pre- and post-WISC-R, the same testers examined 362 subjects to minimize between-tester error. Post-test blood samples were also gathered from those children who gave blood before and agreed to give a second sample.

According to the WISC-R instruction manual, when the WISC-R is repeated a second time a few weeks later, the average test–retest learning is 9 points on non-verbal I.Q. Thus, a 9 point gain was expected after 10–13 weeks. The manual indicates that the standard deviation on non-verbal I.Q. after a few weeks is 4.5 points. With this information, the senior researcher operationally defined "dramatic gains" or "responder" on non-verbal I.Q. as any gain of 15 or more points. The first 9 points of gain could be simple test–retest learning, but an additional 6 points or more would occur by chance only 10 times out of 100. (Six points divided by a standard deviation of 4.5 points equals a Z score of 1.33 or 0.90 units under the normal curve with a one-tenth test.) Conversely, any student who gained less than 7 points in non-verbal I.Q. was defined as a "non-responder".

Table 3. Participation by test

	Number of pre-tests	Number of post-tests	Post-test completion (%)
MAT	612	559	91
EPQ	602	572	93
Raven's Matrices	*	555	90
WISC-R	440	411	67
Blood samples	340	326	53
CTBS	*	296	48
RT/IT	105	75	12

*The Raven Matrices was tested only once at 4 weeks.
296 students completed all portions of the pre- and post-test given over 4 days before the project started and upon completion. $N = 615$.

Secondary measures of validity

Although the WISC-R was the primary indicator of change in non-verbal intelligence, the validity of changes on the WISC-R was examined with two independent measures of performance, the CTBS and RT/IT.

Comprehensive Test of Basic Skills

The CTBS is given in California to all children in the fall and 6 months later in the spring. The test takes 4 full days to give and examines 13 different types of basic skills: vocabulary, spelling, comprehension, reading, language mechanics, language expression, total language, maths comprehension, maths application, total maths, reference skills, science, social studies, as well as a battery score. The fall scores were used as a pre-test and the spring scores as a post-test. The 13 weeks of supplementation was started so that the spring CTBS tests coincided with the last week of supplementation. If the changes in non-verbal I.Q. are real and meaningful, the senior author hypothesized that the group producing the greatest gains in non-verbal I.Q. would also produce the greatest gains on the CTBS.

Reaction time/inspection time

The literature shows that intelligence is highly correlated with many variables that can be measured in milliseconds by a computer that drives a reaction time and inspection time apparatus (Kirby & Nettlebeck, 1989; Frearson & Eysenck, 1986). As a test of the validity of "dramatic" change as indicated by the WISC-R, 75 subjects aged 12–13 years from Mae Hensley were pre- and post-tested on the IT/RT computer by one assistant.

Table 3 shows the number of pre-tests, post-tests, and percent of students completing both for the MAT, EPQ, WISC-R, RT/IT, CTBS, Raven's Matrices, and blood samples.

MAIN RESULTS

A comparison of group means

The Raven Matrices were given after 4 weeks to determine if a significant difference existed between the placebo group and any of the supplement groups after a very short period of supplementation. No significant results were in fact found (Table 4) although all supplementation groups do better than the placebo group.

The results using Raven's Matrices support Nelson *et al.* (1988) finding that a low dose supplement has no significant effect on non-verbal I.Q. after just 1 month. In this study, neither group who received the 50 or 100% formulas produced significantly higher non-verbal I.Q.s than the placebo group after receiving pills for 1 month. However, the group receiving the strongest dose supplement performed significantly better after only 4 weeks of supplementation ($F = 4.81$, $P = 0.029$). The combination of low dose, short time period, and a smaller sample size may have been responsible for Nelson's negative results. In view of these results, there may be a relation between supplementation and non-verbal I.Q. at 4 weeks, but the data are not conclusive.

It is interesting that it is the *highest* dose which gives significant results, while as shown later it is the 100% RDA supplementation which gives the best results. The data suggest that effects are a product of duration of administration multiplied by strength of dose. Thus 4 weeks multiplied

Table 4. Raven Matrices scores after 4 weeks by treatment group

Variable	Mean	SD	N
For entire population	15.61	5.696	489
Placebo	15.1	5.745	130
50% Supplement	15.5	5.609	120
100% Supplement	15.3	5.728	122
200% Supplement	16.7	5.629	117

$F = 1.905$; $P = 0.128$.

Table 5. Change in non-verbal intelligence by treatment group as indicated by the WISC-R

Variable	Mean	SD	N
For entire population	10.52	8.01	410
Placebo	8.9	7.3	100
50% Supplement	10.1	8.9	100
100% Supplement	12.6	7.9	105
200% Supplement	10.4	7.6	105

$F = 3.862$; $P = 0.01$.

by 200% RDA supplementation equals 13 weeks multiplied by 100% RDA supplementation. The data do not prove this hypothesis, of course, but they may suggest an interesting experiment.

No significant differences in verbal intelligence were found on the WISC-R. The 4 groups produced a mean gain of 2 points, exactly what is expected based on the instruction manual. The same finding did not occur for non-verbal intelligence on the WISC-R (see Table 5).

The difference of 3.7 points in gain on non-verbal intelligence between the placebo group and 100% supplement group is statistically significant ($P = 0.01$). In contrast, neither the 50 or 200% group did significantly better than the placebo group. When one controls for compliance by excluding students who participated < 50 days, the 100% supplement group gained 4.4 points more than the placebo group ($P = 0.002$). An analysis of covariance was completed using age, sex, race, school and compliance as potential covariates. None of these variables were significantly correlated with assignment to group. Furthermore, the relationship between the placebo and 100% RDA group held for males and females; whites and non-whites; younger and older students; and in all 4 schools. More specifically, in each comparison the 100% supplement group showed greater gains than the placebo group. Although it was previously mentioned, it is important to remember that a 9 point gain was expected in all 4 groups due to test–retest learning. Thus, the 100% supplement group really gained 4 rather than 13 points. Full data on all sub-tests are given in Table 6.

The students receiving any supplement formula did not out-perform the students receiving placebos or the children reported in the WISC-R manual on the verbal I.Q. subscales or verbal I.Q. In marked contrast, the 105 children receiving the 100% USRDA formula gained more than the 104 subjects in the WISC-R manual on all five non-verbal scales and gained more than the 100 children on placebos on all five non-verbal scales, a 10 to 0 ratio.

The gains in non-verbal intelligence were primarily due to gains in Object Assembly and Coding followed by Picture Arrangement. Block Design produced the smallest amount of the gain of the five non-verbal scales. Only three subscales were significantly different using a 1-tailed test at the 0.05 level: Object Assembly ($F = 2.515$; $P = 0.03$); Coding ($F = 2.234$; $P = 0.04$) and Arithmetic ($F = 4.458$; $P < 0.002$).

The absence of a significant relationship using the MAT was not totally unexpected. Since the test is given in a group setting, the opportunity for students to take the test lightly and rush through it was present. The large standard deviation suggests a number of students did just that. However,

Table 6. Gains on second testing

Scales	Placebo group (100)	50% USRDA (100)	100% USRDA (105)	200% USRDA (105)	All groups (410)	WISC-R norms (104)
Information	0.5	0.4	0.4	0.6	0.5	0.4
Similarities	0.3	0.3	0.5	0.2	0.3	0.7
Arithmetic	0.5	−0.1	0.7	−0.3	0.2	0.5
Vocabulary	0.4	0.1	0.4	0.2	0.3	0.2
Comprehension	0.2	0.3	0.0	−0.2	0.1	0.8
Digit span	0.4	0.7	0.7	0.1	0.5	0.5
Verbal I.Q.	3.0	1.9	2.8	0.4	2.0	3.2
Picture Completion	1.7	1.5	1.5	1.7	1.6	1.2
Picture Arrangement	1.4	1.8	2.1	1.9	1.8	1.7
Block Design	0.8	− 1.2	1.1	0.6	0.9	0.9
Object Assembly	1.2	1.2	2.1	1.3	1.5	1.3
Coding	1.4	1.2	1.9	1.7	1.5	0.9
Non-verbal I.Q.	8.9	10.1	12.6	10.4	10.5	9.2

Note: The column labeled WISC-R norms contains data on 104 children aged 14.5–15.5 years who were retested 3–5 weeks later and published as learning reference norms in the WISC-R instructor manual, p. 33.

Table 7. Change in the Matrix Analogies Test by group

Variable	Mean	SD	N
For entire population	1.4	5.7	558
Placebo	0.9	6.0	147
Half	1.3	5.2	139
Whole	1.8	5.5	136
Twice	1.6	5.9	136

$F = 0.668$; $P = 0.572$.

Table 8. Change in non-verbal intelligence on the WISC-R for four selected groups

Amount of increase	Groups				
	Placebo	50%	100%	200%	
Under 15 points	80 (80%)	68 (68%)	58 (55%)	74 (71%)	280
15+ points	20 (20%)	32 (32%)	47 (45%)	31 (30%)	130
Total	100	100	105	105	410

$\chi^2 = 14.828$; d.f. = 3; $P = 0.002$.

the trend between the 4 groups was the same as found on the non-verbal half of the WISC-R. More specifically, the placebo group did the worst and the 100% RDA group scored the best, making the MAT supportive of the main findings (see Table 7).

As stated in the literature review, a comparison of mean differences may be misleading since no change is expected in most students with dramatic changes appearing in others. A comparison of students who gained less than 15 points with those who gained 15 or more points, shows that the 100% supplement group had 45% of its members gaining 15 or more points. In marked contrast the placebo group had only 20% subjects gaining 15 or more points, a 25% difference in the proportion of responders ($P = 0.002$).

The data clearly support the hypothesis that the mean difference of 3.7 points in non-verbal I.Q. is not a good representative of how much the typical child will gain when nutrition is improved. Approximately one-third of the population produced gains that exceeded the upper limits of the 90th confidence interval, i.e. 14 points. These 130 children gained an average of 20 points in non-verbal I.Q. and produced gains ranging from 15 to 30 points (see Table 8).

The operational definition of "responder", i.e. a gain of 15 or more points, should have produced approx. 10 responders in each of the 4 cohorts of 100 subjects. Yet, the placebo group had 20 responders. An analysis of the placebo responders by school and class showed that 6 came from Mae Hensley, 1 from Riverbank, 3 from Oakdale, and 10 from Ceres High with 9 of the 10 in one class! Closer scrutiny showed that this one class of 70 student athletes had been encouraged to participate too well. Many of them were not only taking the pills they were given in school, but had started taking supplements at home to improve their chances of winning. Fortunately, this pattern did not appear in any other classes or schools, but may explain the inflated number of "placebo" responders.

A further analysis of the 5 scales that make up the non-verbal portion of the WISC-R showed that the unexpected gains in I.Q. were limited to 4 of the 5 scales. The change in Block Design, i.e. 0.93 points, did not exceed what would be expected according to the instructor's manual in normal test–retest situations after 13 weeks. In contrast, unexpected gains in non-verbal I.Q. were found in the 4 remaining non-verbal scales, i.e. picture completion, picture arrangement, object assembly, and coding. In subsequent replications of this study, it would be worthwhile to investigate if the Block Design scale contributes to or subtracts from the gain in non-verbal intelligence scores that appear following supplementation.

CTBS and RT/IT results

A decision was made to exclude from the analysis any students who did not complete all sections of the CBTS in the fall and spring for three reasons. First, those children who did not take all sections of the test were absent and most likely ill. It makes no sense to compare the performance of children who are coming down with or just getting over an illness. Second, any children who were absent during the post-test period were not receiving their supplements and may have had their performance impaired. Third, it is well known that nutrient needs rise during illness and a recovering child is more likely to have deficiencies even if being supplemented. This decision reduced the sample to 296 subjects from 332.

The mean gain in performance for each cohort was calculated and then the 4 cohorts were rank-ordered with a score of 1 being the greatest gain and 4 being the smallest gain. As Table 9 shows, the 100% USRDA group produced the greatest gains in 10 of 13 comparisons and all 8 basic subject areas. In marked contrast, the placebo group ranked last or next to last on all 13

Table 9. Rank order performance on the comprehensive test of basic skills, the WISC-R, the MAT, and Raven Matrices among 4 cohorts

	Placebo group	50% group	100% group	200% group
CTBS ($N = 296$)				
Vocabulary	3	4	1	2
Comprehension	4	3	1	2
Reading	3	4	1	2
Spelling	3	4	1	2
Language	4	2	1	3
Mechanics	4	2	1	3
Expression	4	1	2	3
Maths	4	2	1	3
Comprehension	4	3	1	2
Application	4	1	3	2
Science	3	4	1	2
Social studies	3	4	1	2
Reference skills	3	4	2	1
WISC-R ($N = 410$)	4	3	1	2
MAT ($N = 558$)	4	3	1	2
Ravens ($N = 489$)	4	2	3	1

scales. It is also worth noting that the 200% group tended to do a little better than the 50% group, just as was found in the WISC-R and the MAT.

The CBTS also reports each student's grade equivalent. For example, a student with a score of 8.2 on comprehension means that the child is reading at the 8th grade (or year) and 2 months level. When the test is repeated 6 months later, the same child, if typical, should score 8.8 which would represent a gain of 6 months on comprehension. By calculating the difference between the fall and spring scores, it was possible to contrast the difference in gain between the groups. The magnitude of the difference between the placebo group and the 100% USRDA group was substantial for certain subjects and minimal for others. For example, during the 3 months of supplementation, the 100% RDA group outgained the placebo group by 4 months on Comprehension and Maths Comprehension, 3 months on the total Battery and Reading, and 2 months on Science and total Maths. The other differences were minimal.

Although the magnitude of the differences between the placebo and supplement group were quite striking, only three of them were significant at the 0.05 level, Comprehension, Battery, and Reading. The reason appears to be the magnitude of variation that the individual children create. Replication with a larger cohort that can determine if the difference was coincidental or due to supplementation is clearly warranted.

Of 75 subjects who completed a pre- and post-test with the Reaction Time and Inspection Time apparatus, 26 produced gains of less than 7 points in non-verbal I.Q. on the WISC-R (i.e. they were "non-responders") and 24 produced gains in excess of 14 points (i.e. they were "responders). These two groups were compared to see if any of the predictors of I.Q. that Kirby and Nettlebeck (1989) and Fearrson and Eysenck (1986) discovered could separate the groups. In theory, if the dramatic increases in non-verbal I.Q. were real, then the 24 responders would be expected to produce significantly better scores on the IT/RT apparatus. Kirby found that inspection time (IT), decision time (DT), the standard deviation of the decision time (SDT), movement time (MT), standard deviation of movement time (SMT), and errors (E) produced a canonical correlation of 0.91 ($P < 0.001$) with I.Q. Thus, the decision was made to perform a canonical correlation using the same variables to determine if they could successfully separate the responders and non-responders (see Table 10).

In this table, RT 1 light refers to trials using *simple* reaction time. RT 4 lights refer to choice reaction time trials. RT odd-man-out refers to the technique used by Frearson and Eysenck (1986) in which 3 lights are illuminated simultaneously, 2 close together and 1 some distance away, and the response is made to the odd-man-out. IT refers to trials using inspection time apparatus.

Five of the seven predictors of cohort involve errors. The children who showed the greatest gain in non-verbal I.Q. produced significantly less errors. Subjects who showed no real gain in non-verbal I.Q. made more errors at the post-test. Kirby and others have indicated that a smaller standard deviation in errors for a subject is correlated with higher intelligence. The responders showed a significant reduction in the standard deviation in errors which also supports the

Table 10. 24 Subjects with I.Q. gains of 15+ points vs 26 subjects with gains under 7 points

Change in	RT 1 light 1st trial errors	RT 1 light 2nd trial errors	RT 1 light all trials errors	RT 4 lights all trials decision time
Non-responders	−0.19	−0.58	−0.85	6.5
Responders	0.00	0.13	0.13	29.1
One-tailed test	$F = 3.67$ $P = 0.030$	$F = 6.03$ $P = 0.009$	$F = 5.04$ $P = 0.014$	$F = 5.30$ $P = 0.013$

	RT odd man out difference between decision time +movement time	IT mean errors	IT standard deviation errors
Non-responders	87	−28	129
Responders	25	164	430
One-tailed test	$F = 6.31$ $P = 0.007$	$F = 2.52$ $P = 0.056$	$F = 2.89$ $P = 0.048$

Canonical correlation 0.62; $P < 0.001$.

hypothesis that the gains in non-verbal I.Q. are valid. The amount of time necessary to decide which light to tap was significantly longer among those subjects who did not produce a gain in I.Q. These variables produced a canonical correlation of 0.62 ($P < 0.001$). Thus, the RT/IT demonstrates that the dramatic gains of 15+ points are valid changes in non-verbal I.Q. (Table 1).

Results concerning changes in the EPQ (Eysenck & Eysenck, 1975) are not included in this report as they are not relevant to the question of improvement in abilities; they will be published separately.

DISCUSSION

Discussing our results, it will be clear that little stress can be laid on the results of the Raven Matrices test, as it was only given on one occasion, so that there was no opportunity of comparing each child's performance *after* supplementation with performance *prior* to supplementation. Given the great variance in scores, the non-significant results are perhaps understandable; overall, the results of the WISC-R (non-verbal) Scales are only significant because the analysis is of repeat performances. Even so, significantly better performance of the 200% RDA group supports our major findings.

Similarly, the results of the Matrix Analogies Test cannot be taken too seriously in view of the fact that the children themselves did not appear to collaborate well, so that there is an unacceptably high error involved in the measurement of I.Q. The fact that in spite of this the results follow the pattern set by the WISC-R (non-verbal) tests, i.e. the placebo group having the lowest improvement score, the 100% RDA group having the highest, may to some slight extent support our major conclusions.

The CTBS tests also furnish us with some support, although as pointed out in the discussion of the results, this support is not very strong. Nor would it be expected to be very strong because many of the tests are measures of crystallized rather than fluid ability, although a certain limited amount of fluid ability would seem to come into the account because we are dealing with learning taking place in the limited time period. However, if this were the only support for the theory here tested, one could certainly not rely on this particular set of tests to demonstrate the validity of the hypothesis. Clearly, our decision as to the validity of the hypothesis that vitamin and mineral supplementation can improve performance on tests of fluid ability, but not crystallized ability, rests fairly and squarely on the results of the WISC-R Scale. There the improvement for the placebo group is almost exactly that traditionally found on repetition in random samples, namely 8.9 points of I.Q. This is a value with which all improvements have to be compared. All the supplement groups surpass the placebo group, and the overall statistical comparison between groups is significant at the 1% level. There clearly is a significant improvement, presumably due to supplementation, of non-verbal intelligence as measured by the WISC-R.

No predictions were made at the beginning of the experiment concerning the degree of supplementation which would give optimal results. In these, there was disagreement, some favoring the 200% supplement, others favoring the 100% supplement. As it happens the evidence suggests that the 100% supplement shows the largest difference, namely 3.7 I.Q. points advance on the placebo group. This becomes 4.4 points when children who participated for fewer than 50 days are excluded, given a P value of 0.002. Clearly it is reasonable to exclude these children because they had only received a limited portion of the supplement, and would thus be disadvantaged compared with those who received all the required doses. This is a high level of significance indeed, particularly when we find that age, sex, race, or school failed to account for any of the differences observed.

It is surprising in a sense that the data have come up significantly, because there are a number of considerations which suggest that there are several factors which confuse the situation. We have already mentioned the fact that some children showed only limited compliance, and had far fewer pills than they should have had; this is a type of chance error which would reduce significantly the chances of finding positive results. We have not excluded these non-compliers from the analyses because to do so might have seemed improper, but the point should be borne in mind.

Another important consideration is that our prediction of improvement would only be made of children who are suffering from some kind of dietary deficiency as far as vitamins and minerals are concerned; many, or probably most, of the children tested would not fall into that category. Strictly speaking, we should only have tested children who were deficient, but of course we could not have known at the beginning of the experiment who was and who was not so deficient. Hence our results will be seriously weakend by the presence of large numbers of children whom the theory would not have predicted to have benefitted from the dietary supplementation.

It is for this reason that the analysis described in Table 8 is so important. We predicted that the number of responders (defined as increasing by 15 points or more in non-verbal I.Q. after supplementation) would be significantly more frequently found in the groups receiving the supplements, rather than in the placebo group, and that the percentage they constituted of those groups would be some indication of the number of children suffering some deficiency in vitamins and minerals. We expected the numbers to be small in the placebo group, although we did find, as explained above, that some children did apparently use some form of supplementation independently to improve their athletic performance. However that may be, the results were clearly significant (at the $P = 0.002$ level), and 45% in the 100% RDA group turned out to be responders. It would be idle to then translate this into an estimate of the number of children suffering dietary deficiencies in the population, but it is legitimate to translate these figures into the statement that, in a population such as ours, dietary supplementation improved fluid intelligence estimates by a minimum of 6 points, with an average of 11 points, and a maximum of 21 points. These are improvements which have to be taken seriously as far as practical applications are concerned.

It is hoped to supplement these arguments at a later stage when the analysis of the blood samples has been completed, so that we can verify directly our assumption that responders and non-responders can be differentiated in terms of assessments of their deficiency in minerals and vitamins as disclosed by analysis of their blood samples. Even without doing this, it is clear that the 100% RDA group is superior to the 50% RDA group, and the 200% RDA group, and even more to the placebo group as far as improvement in non-verbal I.Q. is concerned.

The fact that the 50% RDA group did not do as well as the group receiving the 100% RDA formula raises some very important questions that this study cannot answer. First, are the RDAs really adequate? If raising children's daily intake up to 100% of the RDAs with a 50% RDA pill does not improve I.Q. as well as a stronger pill, then are they not too low? Furthermore, if the American RDAs are not adequate, are not the much lower British RDAs even further off?

The fact that the 200% RDA group did not do as well as the 100% RDA group also raises questions. Is the reason that the higher dose supplement did not work as well due to an excess of selected nutrients impairing the digestion, absorption, and utilization of the other nutrients? The data may suggest that "more is not necessarily better", a hypothesis that begs further research. One fact is clear. The three supplement strengths did not produce equal results and greater care should be paid to dose strength in subsequent research, perhaps in connection with duration of administration of the supplements.

The improvements on CTBS scores among children who produced substantial gains in intelligence not only verifies the validity of the changes in I.Q., but also demonstrates the applied impact of raising intelligence on school performance. In light of these results, one may conclude that the increase in performance in New York following a change in their diet policies may have been partially due to nutrition. It is worthwhile to note that New York accomplished the gains without pills, just a well-balanced diet that curtailed sugar, fat, and a few non-essential additives.

An analysis of the blood data is beyond the scope of this report which was limited to testing two primary hypotheses. The question of which nutrients were involved and possible mechanisms will be the subject of a later report. For now it is sufficient to state that a preliminary analysis by the senior author appears to support three findings: (1) the supplements produced significant changes in blood nutrient concentrations for selected vitamins and minerals; (2) "responders" had significantly different changes in blood nutrient concentrations than "non-responders"; and (3) "responders" were more likely to have low blood nutrient concentrations at the pre-test. However, it must be stressed that the blood analysis is preliminary, tentative, and needs to be scrutinized more closely before any final results can be reported.

Our general conclusion, therefore, must be that our general hypothesis has been supported by the results of this study, and that 100% RDA supplementation in minerals and vitamins serves to increase performance on the non-verbal scales of the WISC-R to a significant extent when compared with placebo groups not receiving such supplementation. As pointed out in the introductory chapter to this Symposium, the finding raises more questions than it answers, but it does at least seem to lead to one firm conclusion, as far as the old controversy about cognitive effects of dietary supplementation are concerned. There are such effects, they act in a predictable manner, and they occur in children who normally would be considered as receiving a sufficient diet. Note in this connection that improvements occurred in all the schools sampled, and not only or mainly in the school at the lowest point of the socio-economic continuum. It is to be hoped that these results will lead to a great deal of further research which is necessary in order to resolve some of the questions raised by our result.

Acknowledgements—The authors would like to thank Act Laboratories, Booker Pharmaceutical, Doctor's Data, MSI, Smith Kline Laboratories, and York Laboratories for providing blood assays, supplements, and support personnel. They would also like to acknowledge the Wacker Foundation, the California State University, Stanislaus Foundation, Virgin Atlantic airlines, Leeds Castle, and Federal Express's support services. Most of all, the authors would like to thank the Dietary Research Foundation's Chairman of the Board, Graham Aaronson who realized the importance of the questions raised and donated the funds to conduct this trial.

REFERENCES

Benton, D. & Buts, J. P. (1990). Vitamin mineral supplementation and intelligence. *Lancet*, *335*, 1158–1160.

Benton, D. & Roberts, G. (1988). Effect of vitamin and mineral supplementation on intelligence of a sample of schoolchildren. *Lancet*, *1*, 140–144.

Cherkin, A. (1987). Interaction of nutritional factors with memory processing. In Essman, W. (Ed.), *Nutrition and brain function*. Basel, Switzerland: Karger Press.

Conners, K. (1989). *Feeding the brain: How foods affect children*. New York: Plenum Press.

Cravioto, J. & Delicardie, E. R. (1970). Mental performance in school-age children: Findings after recovery from severe malnutrition. *American Diseases of Childhood*, *120*, 404–410.

Crombie, I. K., Todman, J., McNeill, G., Florey, C. Du V., Menzies, I. & Kennedy, R. A. (1990). Effect of vitamin and mineral supplementation on verbal and non-verbal reasoning of schoolchildren. *Lancet*, *335*, 744–747.

Eysenck, H. J. (1979). *The structure and measurement of intelligence*. New York: Springer.

Eysenck, H. J. & Eysenck, S. B. G. (1975). *Manual: Eysenck Personality Questionnaire*. San Diego: Edits. London: Hodder & Stoughton.

Frearson, W. & Eysenck, H. J. (1986). Intelligence, reaction and a new "odd-man-out" R.T. paradigm. *Personality and Individual Differences*, *7*, 807–817.

Kirby, N. H. & Nettlebeck, T. (1989). Reaction time and inspection time as measures of intellectual ability. *Personality and Individual Differences*, *10*, 11–14.

Matarazzo, J. D. (1972). *Wechsler's measurement and appraisal of adult intelligence*. Baltimore: Williams & Wilkins.

Naglieri, J. A. (1985). *Matrix Analogies Test—short form*. Columbus, Ohio: Merrill.

Nelson, M., Naismith, D. J., Burley, V. J. & Gatenby, S. J. (1988). Can children's intelligence be increased by vitamin and mineral supplements? *Lancet*, 6 August.

Rush, D., Stein, Z. & Susser, M. (1980). A randomized controlled trial of prenatal nutritional supplementation in New York City. *Pediatrics*, *65*, 683–697.

Sandstead, H. (1986). Nutrition and brain function: Trace elements. *Nutrition Review*, *44*, 37–41.

Schoenthaler, S. (1991). *Improve your child's I.Q. and behavior*. London: BBC Books.

Schoenthaler, S. J., Doraz, W. E. & Wakefield, J. A. Jr (1986a). The impact of a low full additive and sucrose diet on academic performance in 803 New York City public schools. *International Journal of Biosocial Research*, *8*, 185–195.

Schoenthaler, S. J., Doraz, W. E. & Wakefield, J. A. Jr (1986b). The testing of various hypotheses as explanation for the gains in national standardized academic test scores in the 1978–1983 New York City nutrition policy modification project. *International Journal of Biosocial Research*, *8*, 196–203.

Underwood, B. (1986). Evaluating the nutritional status of individuals: a critique of approaches. *Nutrition Review*, *44*, 213–224.

Yudkin, J. (1988). Letter to the editor. *Lancet*, *1*, 407.