THE FINDINGS AND CONCLUSIONS IN THIS REPORT HAVE NOT BEEN FORMALLY DISSEMINATED BY FDA AND SHOULD NOT BE CONSTRUED TO REPRESENT ANY AGENCY DETERMINATION OR POLICY.

DEPARTMENT OF HEALTH & HUMAN SERVICES Public Health Service
Food and Drug Administration

Memorandum

Date: September 1, 2010
Subject: Citizen Petition from the Center for Science in the Public Interest (CSPI) requesting the revocation of the color additive approvals of eight synthetic dyes for use in food – **Interim Toxicology Review Memorandum**

Docket No. FDA-2008-P-0349
Correspondence No. 84566

Center for Science in the Public Interest
1875 Connecticut Ave., N.W., Suite 300
Washington, DC 20009

CSPI has submitted a petition requesting that FDA revoke approvals for eight certified color additives for use in foods: FD&C Blue 1 and 2, FD&C Green 3, Orange B, FD&C Red 3, FD&C Red 40, and FD&C Yellow 5 and 6. In support of this request, CSPI cites findings from a study that was conducted by the University of Southampton on the effect of certain mixtures of color additives in children; this study was published in the *Lancet* in 2007[1]. CSPI also references a number of publications, including a 2004 meta-analysis by Schab and Trinh[2] of previously conducted clinical trials that investigated the proposed relationship between consumption of artificial food colors and behavioral changes in children diagnosed with hyperactivity. This interim review memorandum along with its four attachments presents FDA's detailed in-depth review and analyses of these study publications, and other relevant published information pertaining to color additives and their potential to cause adverse neurobehavior in children. The review was conducted by the Toxicology group of FDA's Center for Food Safety and Applied Nutrition, Office of Food Additive Safety, Division of Petition Review (DPR).

In our evaluation of this petition request FDA has conducted a thorough review of the Southampton Study from the perspective of its design, results and conclusions. The FDA also has reviewed the other publications included in this petition, including the 2004 Schab and Trinh meta-analysis publication. In addition, FDA conducted a comprehensive search, review and analysis of the scientific literature on color additives and their potential relationship to neurobehavioral disorders in children. Both ToxLine and PubMed databases were searched to identify articles related to the effect of consumption of color additives on children's behavior, with a focus on clinical trials. The criteria used for the review and the search terms are detailed

[1] McCann, D. et al. 2007. Food additives and hyperactive behaviour in 3-year old and 8/9-year old children in the community: a randomized, double-blinded, placebo-controlled trial. *The Lancet*, November, 2007, Volume 370, Issue 9598, pg. 1560-15-67.
[2] Schab, DW and Trinh, NT. 2004. Do artificial food colors promote hyperactivity in children with hyperactive syndrome? A meta-analysis of double-blind placebo-controlled trials. *J. Dev. Behav. Pediatr.* 25(6): 423-34.

in attachment 1. Articles deemed most relevant were selected for in-depth review based on the abstracts retrieved in the search.

Our review reports of the 2007 Southampton study and the 2004 Schab and Trinh meta-analysis are included as attachments to this cover memorandum (attachments 2 and 3). Our review and analysis of other published scientific literature is included as attachment 4, which incorporates the overview report, reviews and critiques of 33 of the most relevant clinical trials identified from the literature search, a tabular summary of these clinical trials, and a bibliography. All of these reviews were conducted by an expert neurotoxicologist through an Interagency Agreement with Oak Ridge National Laboratories (ORNL). The draft ORNL review reports were reviewed by an OFAS toxicologist and have been accepted as final.

In summary, based on the data reviewed in these publications FDA concludes that a causal relationship between exposure to color additives and hyperactivity in children in the general population has not been established. However, for certain susceptible children with ADHD and other problem behaviors, the data suggest that their condition may be exacerbated by exposure to a number of substances in food, including, but not limited to, artificial food colors. Findings from relevant clinical trials indicate that the effects on their behavior appear to be due to a unique intolerance to these substances and not to any inherent neurotoxic properties.

ATTACHMENT 1: Review Criteria and Search Terms Used in ORNL Literature Review
ATTACHMENT 2: Finalized Review of 2007 Southampton Study – ORNL
ATTACHMENT 3: Finalized Review of 2004 Schab and Trinh Publication – ORNL
ATTACHMENT 4: Finalized Review and Analysis of the Literature – ORNL

THE FINDINGS AND CONCLUSIONS IN THIS REPORT HAVE NOT BEEN FORMALLY DISSEMINATED BY FDA AND SHOULD NOT BE CONSTRUED TO REPRESENT ANY AGENCY DETERMINATION OR POLICY.

Review Criteria and Search Terms Used in ORNL Literature Review

The FDA submitted a work assignment (WA 2008 – 31) through a standing Interagency Agreement with Oak Ridge National Laboratories (ORNL) to conduct a search of the literature for publications that addressed the possible involvement of food ingredients in neurobehavioral issues in children. The FDA also requested that an in-depth review and analysis of all selected publications be conducted. The work assignment requested that;

- ORNL conduct a thorough search of the literature from 1982 to the present for peer-reviewed publications dealing with consumption of additives in food and their possible correlation with hyperactivity and behavioral changes. The search should include both animal and human studies. TOXLINE and PUBMED were to be searched in combined searches for the years 1982-2008. The following terms were to be used:
  Food additives hyperactivity, Food additives autistic, Food additives psychomotor, Food additives attention deficit, Food additives neurotoxic, Food additives behavior neurologic*, Food coloring hyperactivity, Food coloring autistic, Food coloring psychomotor, Food dyes hyperactivity, Food dyes autistic, Food dyes psychomotor, Food dyes attention deficit, Food dyes neurotoxic, Food dyes neurologic*, Food dyes behavior.

- All recovered references and abstracts were to be submitted to the ORNL reviewer for his selection of publications for in-depth review. If the reviewer rejected any article as not appropriate for review, he was to write a brief statement giving reasons for this rejection. The reviewer was to summarize studies that demonstrated a causal relationship between food additives and behavioral and/or hyperkinetic changes, and give his assessment as to the overall quality of the study and whether he felt that the study results supported a correlation. In addition, the reviewer was to note any deficiencies in the studies he reviewed and make recommendations as to how these deficiencies should be addressed in future studies. The reviewer was also to review any clinically-relevant publication that was included as part of the CSPI citizen petition.

- Once the reviewer completed his review he was to submit a completed report to FDA detailing his findings and giving an overall conclusion on his opinion regarding the key question. The FDA would conduct a secondary review of this report and submit its comments and requested changes to the reviewer for incorporation in the completed report.

THE FINDINGS AND CONCLUSIONS IN THIS REPORT HAVE NOT BEEN FORMALLY DISSEMINATED BY FDA AND SHOULD NOT BE CONSTRUED TO REPRESENT ANY AGENCY DETERMINATION OR POLICY.

STUDY EVALUATION REPORT

Chronic and Acute Effects of Artificial Colourings and Preservatives on Children's Behaviour

STUDY TYPE: Clinical Study
University of Southampton School of Psychology
Project Code: T07040

Prepared for
Center for Food Safety and Applied Nutrition
Office of Food Additive Safety
U.S. Food and Drug Administration
College Park, MD 20740-3835
Task Order (Assignment) No. 2008-32

Prepared by

Toxicology and Hazard Assessment Group
Environmental Sciences Division
Oak Ridge National Laboratory
Oak Ridge, TN 37831

Primary Reviewer: Tom Sobotka, Ph.D.,	Signature: Date: 6-29-10
Secondary Reviewers: Carol Wood, Ph.D., D.A.B.T.	Signature: Date: JUN 29 2010
Robert H. Ross, M.S., Group Leader	Signature: Date: JUN 29 2010
Quality Assurance: Lee Ann Wilson, M.A.	Signature: Date: JUN 29 2010

Oak Ridge National Laboratory managed and operated by UT-Battelle, LLC., for the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.

TABLE OF CONTENTS

LIST OF TABLES ***4***

Evaluation Report ***5***

I. Executive Summary ***5***

II. Introduction – Study Objectives ***7***

III. Study Design/Methods ***7***

A. Primary Study **7**
1. Participants 7
2. Primary Study Design 7
3. Challenge Treatments 8
4. Challenge Protocol 9
5. Behavioral Measures 9
6. Data Analysis Methods 10

B. Additional Experimental Design Procedures to Address Salient Correlative Secondary Research Questions **11**
1. Effects of the AFCA challenge mixes on component behavior measures of GHA 11
2. Behavioral effects and metabolic mediators following acute challenge with AFCA 11
3. Modulation of AFCA behavioral effects by genetic polymorphisms 12

C. Reviewer Comments Regarding Study Design/Methods **13**
1. ***Questionable accuracy of teacher pre-study behavioral ratings recalling behavior over a period of 6 months*** 13
2. ***Questionable use and accuracy of pre-study behavioral ratings by participating parents*** 13
3. ***Disproportionate increases were made in levels of artificial colors in the mixes (A and B) used for the 8/9YO children relative to the mixes (A and B) used for the 3/4YO children, and no adjustments were made to levels of sodium benzoate in either Mix A or B for either age group of children*** 13
4. ***Composition of placebo is unclear*** 13
5. ***Parents were not asked about the blinding*** 14
6. ***Timing of juice consumption at home*** 14
7. ***Precision of recall data in conduct of weekly parent/teacher ratings and availability of standard ratings*** 14
8. ***Reliability of modified Classroom Observation Code (COC) used in 3/4YO children is unclear*** 14
9. ***Questionable adequacy of duration used for classroom observation*** 14
10. ***Model 1 is not useable for analysis of data*** 14
11. ***In the "acute challenge phase" of this study the disaggregated behavior measures comprising the Hyperactivity Index (HI) were not analyzed separately*** 15
12. ***The Mixed Model analyses of the "acute challenge" data did not appear to control for potential confounding factors*** 15
13. ***No description of materials or methods was provided for the genotyping analyses*** 15

IV. Study Investigators' Reported Findings ***15***

A. Reported Study Findings Relative To Primary Research Question: Do Mixtures of Certain Artificial Food Colors and Sodium Benzoate Increase the Mean Level of Hyperactive Behavior (*Hyperactivity Profile Behaviors*) In Children From the General Population? **15**
1. Study Sample and Background Characteristics 15
2. Juice consumption and dietary infractions 16
3. Effects of Challenge 16

B. Reviewer Comments Regarding Reported Study Findings for Primary Research Question **18**
1. ***Limited information regarding CPT performance decrements of 8/9YO children over study period*** 18
2. ***Problematic aspects of the data analyses*** 18

3. *There is no clarification of the behavioral changes in this study as necessarily "adverse" (p.22) or "deleterious" (p.26)* 19

C. Reported Study Findings Regarding Secondary Research Questions 19

1. (Secondary Research Question) Effects of the AFCA challenge mixes on component behavior measures of GHA 19

2. *Reviewer Comments:* Effects of the AFCA challenge mixes on component behavior measures of GHA 21

(a) *Complementary changes across disaggregated behavior measures can provide reliable confirmation of significant treatment effects.* 21

(b) *Use of the three sample groups of children for analysis of the component (disaggregated) behavior measures results in variable statistical outcomes but relevance of this to data interpretation are not discussed* 21

3. (Secondary Research Question) Behavioral effects and metabolic mediators following acute challenge with AFCA – Study Findings and Interpretation 22

4. *Reviewer Comments*: Behavioral effects and metabolic mediators following acute challenge with AFCA – Study Findings 22

(a) *Statistical analyses of data in the "acute challenge" component study did not control for confounding variables.* 22

(b) *Rationale for explaining why treatment effects in the main study were detected only in parental ratings was based on inaccurate interpretation of acute study findings.* 23

5. (Secondary Research Question) Modulation of AFCA behavioral effects by genetic polymorphisms – Study Findings 23

6. *Reviewer Comments*: Modulation of AFCA behavioral effects by genetic 26

polymorphisms – Study Findings 26

(a) *Variability in finding of significance for main effect of challenge (Mix A and Mix B) on behavioral responses (GHA) with repeated Mixed Models analyses for different genotypes raises questions about replicability of the Mixed Models analyses and uncertainties about the findings of significant treatment effects in the primary study.* 26

(b) *The moderating influence specifically of theThr105Ile/present genotype decreasing the behavioral effects of Mix A in 3/4 YO children appears contradictory to the main effect of genotype on GHA levels.* 27

(c) *The findings from the analysis of the genotype component study data which involved children from the general population were inappropriately extrapolated to ADHD children.* 27

7. Additional Secondary Questions – Study Findings 27

(a) *Consistency in response between Mix A and Mix B.* 27

(b) *Dose-response relationship.* 28

(c) *Difference between the 3/4 year old and 8/9 year old children in changes in GHA over study period.* 28

V. Study Investigators' Overall Conclusions of Study 28

VI. Reviewer Conclusions 29

VII. REVIEWER NOTED STUDY STRENGTHS 32

VIII. REVIEWER NOTED STUDY WEAKNESSES 32

A. Procedural Weaknesses 33

B. Weaknesses Regarding Data Analyses and Interpretation of Study Findings 34

IX. Applicability to Assess Risk or to Support Regulatory Action 37

References 38

LIST OF TABLES

TABLE 1. Composition of Challenge Material 8
TABLE 2. Numbers of 3/4-Year Old and 8/9-Year Old Children in Study 15
TABLE 3. Unadjusted Mean GHA Scores* for 3/4YO and 8/9YO Children by Challenge Type 16
TABLE 4. General Linear Mixed Models estimates of GHA effect size for challenge Mix A versus Placebo and challenge Mix B versus Placebo 17
TABLE 5. Effect of sizes of Mix A versus Placebo and Mix B versus Placebo for each disaggregated behavior measure 20
TABLE 6. Mixed Models analyses of interactions between Genotype and Behavioral Responses (GHA) to Additive Mix A and Mix B versus Placebo in the ≥85% consumption sub-sample of 3/4 YO Children with All Potential Confounds Controlled 24
TABLE 7. Mixed Models analyses of interactions between Genotype and Behavioral Responses (GHA) to Additive Mix A and Mix B versus Placebo in the ≥85% consumption sub-sample of 8/9 YO Children with All Potential Confounds Controlled 25

Evaluation Report

Project Title: Chronic and acute effects of artificial colourings and preservatives on children's behaviour

Principle Investigators: Jim Stevenson (PI)
Edmund Sonuga-Barke (co-PI)
John Warner (co-PI)

Contractor: School of Psychology, University of Southampton (England)

Study Dates: September 1, 2004 to February 28, 2007

Study Technical Report Date: Submitted June 18, 2007; Revised September 10, 2007

The format for this evaluation of the Southampton study on artificial colors/preservatives and children's behavior includes an Executive Summary followed by several sections summarizing information provided by the study investigators regarding: the study objectives; the study design and methods; the reported study findings (addressing the primary and secondary research questions); the investigators' overall conclusions; and the reviewer conclusions on the study. Each of these sections will include "reviewer" comments, specifically addressing the information in that section. The final three sections of this report will address: the strengths of the study, study weaknesses, and the applicability of the study findings for assessing risk or supporting regulatory action.

I. Executive Summary

The principle investigators pointed out that there is a longstanding suggestion, initiated by Ben Feingold (1975) more than 30 years ago, that artificial food colors and additives (AFCA), including preservatives, have detrimental effects on children, inducing an adverse level of overactive, impulsive and inattentive behaviors, i.e. "hyperactivity". Children who show this behavior pattern to a marked degree are also likely to be diagnosed with attention deficit hyperactivity disorder (ADHD). Although similar types of behaviors also occur among the general population of children, it is important to emphasize the distinction between the mild nature of these behaviors in the normal population (which is the focus of the present study) versus the severe, persistent and disruptive nature of these types of behaviors which characterize the abnormalities of ADHD. Earlier studies have failed to confirm the suggested causative association between AFCA and ADHD in children, although a recent meta-analysis of double-blind, placebo-controlled clinical studies reported a significant effect of AFCA on the behavior of children with ADHD. Whether AFCAs have a similar effect in the general population has not been conclusively demonstrated. The principle investigators of the present study conducted a previous study which provided some suggestive evidence of effects on *hyperactivity profile behaviors* based on parental ratings for 3 year old children from the general population in response to a mix of AFCA, but these findings were not replicated by concomitant clinical behavioral assessments.

The <u>primary hypothesis</u> tested in the present study was that mixtures of certain artificial food

colors and sodium benzoate (AFCA) increase the mean level of *hyperactivity profile behaviors* in two age groups of children (3 to 4 years old and 8 to 9 years old) from the general population. This study was designed in part to replicate the principle investigators' previous findings for 3 to 4 year old children and to extend those findings to test whether effects could be identified in 8 to 9 year old children from the general population. The study design was a double-blind, placebo-controlled, cross-over challenge with each treatment administered in fruit juice drinks daily for 1 week. Two mixes of artificial food colors and sodium benzoate were tested using several measures of *hyperactivity profile behaviors*, including parent ratings, teacher ratings and classroom observations, as well as computerized continuous performance testing in the older children. To measure individual differences in *hyperactivity profile behaviors* using these different sources of behavior measures, a Global Hyperactivity Aggregate (GHA) score was calculated as the unweighted composite aggregate of the standardized weekly parent rating, teacher rating, classroom observation, and continuous performance task (CPT) z-scores. A high GHA level indicated more *hyperactivity profile behaviors*.

Based on the analysis of the *whole sample* of children (considered the primary analysis), the findings from the primary study were that challenge with Mix A of artificial food colors and benzoate preservatives elicited statistically significant increases in GHA levels (greater *hyperactivity profile behaviors*) of 3/4 year old children and that challenge with Mix B elicited statistically significant increases in GHA levels (greater *hyperactivity profile behaviors*) of 8/9 year old children. Additional analyses were conducted to assess the effects of both challenge mixes on each individual (disaggregated) component behavior measure. While the latter analyses indicated that the parental rating was the major contributor for the primary effects of Mix A on GHA levels in 3/4 year old children, the parental rating measure alone did not show a statistically significant effect for Mix A challenge. In the 8/9 year old children the parental rating and continuous performance task (CPT) scores were the major contributors for the primary effects of Mix B on GHA levels but only the parental rating showed a statistically significant treatment effect for Mix B. In neither age group of children were there any significant treatment related changes detected by teacher ratings or classroom observations. The principle investigators' previous study with 3 year old children found a significant behavioral effect of Mix A based specifically on parental ratings. This previously reported effect of Mix A, based specifically on parental ratings was not replicated in the present study in the 3/4 year old children, although the Mix A challenge did elicit a statistically significant increase in the overall GHA levels (aggregate *hyperactivity profile behaviors*).

Overall, the primary study findings are suggestive of low level behavioral effects of a one week exposure to AFCAs on behavior in 3/4 year old children (Mix A) and 8/9 year old children (Mix B), based solely on parental ratings. However, due to the absence of confirmation of the parentally identified treatment effects by any other behavior measures together with the concerns about the data analyses and various procedural weaknesses of this study, it is the reviewers' opinion that there is questionable confidence in the reliability and biological relevance of the primary findings from this study. One particular procedural weakness relevant to regulatory application was the use of chemical mixtures as challenge materials which basically precludes identifying which specific compound(s) within the mixtures might be responsible for any treatment related effects. Consequently, there would be little, if any, utility of these findings to assess risk or to support regulatory decisions for specific compounds.

II. Introduction – Study Objectives

The primary hypothesis tested in the present study was that mixtures of certain artificial food colors and sodium benzoate (AFCA) increase the mean level of *hyperactivity profile behaviors* in two age groups of children (3 to 4 years old and 8 to 9 years old) from the general population. The study design was a double-blind, placebo-controlled, cross-over challenge with each treatment administered in fruit juice drinks daily for 1 week. Two mixes of artificial food colors and sodium benzoate were tested using several measures of *hyperactivity profile behaviors*, including parent ratings, teacher ratings and classroom observations, as well as computerized continuous performance testing in the older children. The more salient correlative secondary questions addressed in this study included:

- whether concordant treatment related effects are seen in teacher ratings, direct observations and test performance, as well as in parental ratings;
- whether behavioral and metabolic changes in children are apparent following acute challenge with a single dose of a food color/sodium benzoate mixture; and
- whether selected genetic differences (polymorphisms) modulate the behavioral effects of AFCA treatment.

III. Study Design/Methods

A. Primary Study

1. Participants

Two age groups of children in Southampton, UK were recruited for participation in this study, one group of 3-4 year olds (3/4YO) and one group of 8-9 year olds (8/9YO). The study sample for the 3/4YO group (n=153 with 79 males/74 females) was drawn from a general population of children registered in 'early-years settings' (EYS) (nurseries, day nurseries, preschool groups, playgroups). The sample for the 8/9YO group (n=144 with 75 males/69 females) was drawn from children attending primary and junior schools. Schools (nine participating) and EYS (26 participating) were selected to reflect the full range of socio-economic background of children in the area. Behavioral screening data from a hyperactivity questionnaire (ADHD Rating Scale-IV/Teacher version) rated prior to the start of the study by teachers for all 3-year-old and 8/9-year old children in participating schools and EYS to indicate the frequency of inattention and hyperactive behaviors over the previous 6 month period also showed that behavioral ratings for the study samples are representative of those for children of the same age in the participating schools and EYS [see *Reviewer Comment 1,* below]. At an initial home visit, written informed consent was obtained from parents who agreed to participate in the study. During the home visits, a report was obtained of each child's pre-study diet based on 24-hour recall by the parents to assess the levels of foods containing additives consumed by the children in the previous 24 hour period. Prior to the start of the study the participating parents completed a behavioral questionnaire (ADHD Rating Scale-IV/Home-Parent version) to indicate the frequency of occurrence of inattention and hyperactive behaviors over the past 6 months [see *Reviewer Comment 2,* below].

2. Primary Study Design

The primary study design and challenge protocols were similar for both age groups of children (3/4YO and 8/9YO). Testing of the two age groups was conducted

consecutively with the 8/9YO group tested first. The experimental design was a double-blind within-subject crossover with two active treatments (Mix A and Mix B) and a placebo administered daily in a fruit-juice drink.

3. Challenge Treatments

The compositions of the two active Mixes for each age group are shown below in Table 1.

TABLE 1. Composition of Challenge Material

	Mix A		Mix B	
3/4 Year Olds	Sunset yellow	5.0 mg	Sunset yellow	7.5 mg
	Carmoisine	2.5 mg	Carmoisine	7.5 mg
	Tartrazine	7.5 mg	Quinoline yellow	7.5 mg
	Ponceau 4R	5.0 mg	Allura red AC	7.5 mg
	Total Colors	***20 mg***	***Total Colors***	***30 mg***
	Sodium benzoate	45.0 mg	Sodium benzoate	45.0 mg
8/9 Year Olds	Sunset yellow	6.25 mg	Sunset yellow	15.6 mg
	Carmoisine	3.12 mg	Carmoisine	15.6 mg
	Tartrazine	9.36 mg	Quinoline yellow	15.6 mg
	Ponceau 4R	6.25 mg	Allura red AC	15.6 mg
	Total Colors	***24.98 mg***	***Total Colors***	***62.4 mg***
	Sodium benzoate	45.0 mg	Sodium benzoate	45.0 mg

For the 3/4YO children the two active mixes (A and B) differed from each other in quantity of color additives and the specific color additives used. Mix A was similar to the active challenge used in the earlier study of 3-year old children by Bateman et al (2004) and Mix B was selected to represent the current average daily consumption of food additives by 3-year olds in the UK. Similarly, for the 8/9YO children the two active mixes (A and B) differed from each other in quantity of color additives and the specific color additives used. Mixes A and B used for the 8/9YO children had higher levels of color additives than the mixes used for the 3/4YO children, in order to account for the increased amount of food typically consumed by older children. However, both Mixes A and B in both the 3/4YO and 8/9YO groups included the same amount of sodium benzoate, a food preservative (see *Reviewer Comment 3,* below).

The placebo drink consisted of a mixture of fruit juices. The specific composition of the placebo drink and the variety of fruit juices used were not described [see *Review Comment 4,* below]. The only difference in the composition of the placebo and active mixes was the presence of the AFCA in the active mix with some variation in the proportions of the fruit juices to ensure matching color and taste for the placebo and active drinks. A masked testing by two independent panels of 20 young adults showed that the active and placebo juice drinks could not be differentiated based on look and taste. During the course of the study, the study administrator assigned the

challenge sequence (Mix A, Mix B and placebo) for each child. The child's family and the research team were masked to the challenge sequence. Identical sealed bottles of juice drinks were delivered to homes every week by the research team [see *Reviewer Comment 5*, below]. The juice was kept in a refrigerator and consumed at home either prior to the child's session in the EYS/school or after return from the EYS/school; the 8/9YO children consumed the juice mainly after returning from school (p. 168) [see *Review Comment 6,* below]. Any bottles with unconsumed juice were returned to the study office. Parents completed a daily diary of juice consumption and monitored compliance with the diet by recording dietary infractions ('mistake events'), when a child consumed a portion of food containing the artificial colors or sodium benzoate.

4. Challenge Protocol

After one week on the normal diet (week 0: baseline), the challenge protocol was conducted over a period of six weeks throughout which the artificial colors and sodium benzoate, to be used in the active challenges, were removed from the children's diet. During the first week of additive (colors/benzoate) withdrawal from the diet, all children received placebo drinks (week 1: withdrawal with placebo); during weeks 2, 4, and 6, each child was scheduled to a randomized set of two active challenge weeks and one placebo challenge week (week 2, 4 or 6: challenge) in which the children were given the appropriate challenge drink daily for seven days; and in weeks 3 and 5 all children received placebo drinks (week 3 or 5: washout with placebo). The 3/4YO children were given daily drinks of 300 mL/day and the 8/9YO children were given daily drinks of 625 mL/day. At the beginning of the study, the study administrator assigned each child using a random number generator to one of six possible sequences of receiving the placebo, active Mix A, or active Mix B challenges across weeks 2, 4, and 6.

5. Behavioral Measures

Behavioral Screening: As noted previously, behavioral screening, using the ADHD Rating Scale-IV questionnaires (teacher and parent versions), was conducted one time prior to Baseline: Week 0 by teachers for all children of the target age within participating schools and EYS (early year settings) and by participating parents to indicate frequency of specific inattention and hyperactive behaviors over the previous 6 months [see *Reviewer Comments 1 and 2*, below].

Weekly Behavior Measures: During the primary study, three measures of behavior were used to assess treatment effects for the 3/4YO children, with an additional fourth measure for the 8/9YO children. (1) Teacher ratings – the abbreviated ADHD Rating Scale-IV(teacher version) was completed by teachers once each week to indicate the frequency of inattention and hyperactive behaviors displayed over the past week for each week of the study (Week 0: baseline to Week 6). (2) Parent ratings – Parents of 3/4YO children used the abbreviated Weiss-Werry-Peters (WWP) hyperactivity scale and parents of 8/9YO children used an abbreviated ADHD Rating Scale-IV (parent version) to rate changes in their children's behavior over the previous week for each week of the study (Week 0: baseline to Week 6) [see *Reviewer comment 7*, below]. (3) Classroom Observation Code (COC) – The COC assesses the occurrence of mutually exclusive behaviors during structured didactic teaching and during periods of independent work under teacher supervision. In developing this measure,

behaviors had been selected to indicate components of ADHD that are shown in the classroom; for 3/4YO children, who in the UK have little structured didactic teaching and are not normally given "tasks" but allowed to choose a range of activities, the observation code was adjusted accordingly. Each child was observed by an independent observer (psychology graduate) for a total of 24 minutes each week of the study in three 8-minute observation sessions from which a total mean weekly score was derived [see *Reviewer comments 8 and 9*, below]. (4) Conners' Continuous Performance Test II (CPTII) - A fourth behavioral measure for 8/9YO children was the CPTII, a computer based test of 14 minutes duration using response to visual stimuli to assess attention and the response inhibition component of executive control brain function. The CPT was administered weekly to the 8/9YO children only. The subject is presented with 18 blocks of 20 trials each (total 360 trials) and tested with three different inter-stimulus intervals (ISI) of 1, 2 or 4 seconds. This CPT is a "not X" task, requiring the subject to press a computer key immediately in response to all letter presentations other than the letter "X". Four test measures (standard error of reaction time, % commission errors, signal detection index - d', and response bias – β) were used to derive a weekly CPT aggregate score. These measures have been shown to be highly correlated with the parent ADHD rating scale measure.

6. Data Analysis Methods

Global Hyperactivity Aggregate (GHA): A Global Hyperactivity Aggregate (GHA) score was developed to measure individual differences in hyperactivity (*hyperactivity profile behaviors*) using different sources of behavior measures (i.e., teacher ratings, parent ratings, classroom observations, and a computerized CPT test). To calculate the GHA, weekly scores from the Parent, Teacher, COC and CPT measures for each child were standardized to time 0 at Baseline (T0; aka Week 0) for the same measure, as:

weekly standardized (z) aggregate score = (score X – mean X at T0)/ SD at T0

The GHA was then calculated as the unweighted composite aggregate of the weekly Parent, Teacher, COC, and CPT z-scores. The GHA was calculated only when at least 3 of the different behavior z-scores were present for any week. The absolute value of the GHA score indicates the relative change in hyperactivity compared to baseline (the higher or more positive the number, the greater the level of hyperactivity behavior compared with baseline and the lower or more negative the number, the lower the levels of hyperactivity relative to baseline).

Statistical Analysis: Although the study designs for the two age groups (3/4YO and 8/9YO) were similar, the difference in composition of the GHA (the additional CPT behavior measure was included in testing of 8/9YO children only), and in the dose and composition of the AFCA mix used, precluded joint analysis of the data from the two age groups. Therefore, for analysis purposes the studies for the two age groups were treated as parallel but independent.

Linear mixed-model methods in Statistical Package for the Social Sciences (SPSS) were used to analyze data. Two models were tested separately for each age group for the effects of treatment on GHA in challenge weeks. Model 1 used the challenge type alone as a fixed effect testing for Mix A versus placebo and Mix B versus placebo. In Model 2, in addition to challenge type, the effects of the following potential confounding factors were controlled: week during study, sex, GHA in baseline week

(Week 0), number of additives in pretrial diet, maternal educational level, and social class [see *Reviewer Comment 10,* below]. The study was powered to detect differences between the active and placebo periods and, accordingly, the effects of Mix A and Mix B were compared with that of placebo.

The analyses were replicated for three subgroups of the study population in each age group: the *whole sample*, a high consumption sub-sample group *(≥85% consumption* of drinks in any challenge week), and a *complete case* sub-sample (consuming >85% of drinks in all challenge weeks and no missing GHA score). The latter two sub-sample groups were included to determine whether non-compliance (not consuming all of the scheduled drinks) and the method of handling missing data affected the pattern of results.

B. Additional Experimental Design Procedures to Address Salient Correlative Secondary Research Questions

1. <u>Effects of the AFCA challenge mixes on component behavior measures of GHA</u>
The present study was designed with a global aggregated measure of hyperactivity (GHA), combining the behavior measures (teacher ratings, parent ratings, classroom observations, and a computerized test), as the primary outcome measure. Treatment effects were calculated as changes in GHA in response to the challenge mixtures compared with placebo. In a previous study of the behavioral effects of AFCAs in 3-year old children (using a challenge equivalent to Mix A in the present study), uncertainty in the interpretation of the results occurred due to the finding that significant effects of the Mix A challenge were detected only by parental reports and were not confirmed by clinical behavioral testing (Bateman et al, 2004). In an effort to determine whether the previous findings were replicable and to extend those findings, additional analyses of the behavior data for the 3/4YO and 8/9YO children in the present study were conducted to address the correlative secondary research question of whether challenge related effects are seen across the different sources of behavior measures in addition to the parent ratings. For this purpose, analyses of the behavioral response of additive challenge versus placebo for each individual behavior measure (*disaggregated measures* of teacher ratings, parent ratings, classroom/playroom observation scores, and CPT) were performed and presented in the final report as a secondary outcome measure.

2. <u>Behavioral effects and metabolic mediators following acute challenge with AFCA</u>
A second phase of this study was designed as a "proof of principle" component to address the correlative secondary question of whether it is possible to demonstrate short term changes in behavior immediately after single dose acute challenge with a mix of AFCAs and to explore their relationship to metabolic factors that may mediate such responses. Participants in this study were enlisted from the 8/9YO boys who participated in the primary phase of the challenge study. Two groups of 15 boys each, who did or did not exhibit a behavioral response to Mix B in the primary challenge study, were identified. Using each child's GHA scores from the primary study, a GHA difference score was calculated by subtracting the Placebo score from the Mix B score. Higher difference scores reflected a more negative behavioral response to Mix B. The GHA difference scores were then ranked and children with scores ≥ 75th percentile were classed as "*responders*" and those with scores ≤25th

percentile as "*non-responders*". Parents of children were asked to resume the reduced additive diet that was followed in the primary phase of this study for a period of 24 hours prior to initiation of the acute challenge phase of study. The Acute Challenge was conducted at the Southampton School of Psychology in two 2.5 hour sessions approximately one week apart. In each session children were given in a random sequence a single dose of either the active challenge (Mix B with amounts of additives equivalent to the daily total administered in the primary study) or the placebo challenge, each administered by capsule. The children's response inhibition and attention were measured using the Conners' Continuous Performance Test II (CPT) twice, approximately 30 minutes prior to and 30 minutes after challenge. While completing the CPT, children's behavior was also monitored by independent observers using the seven item Hillside Behavior Rating Scale. In addition to the behavioral measures, urine samples were collected to test for histamine which is thought to mediate a pharmacological effect of additives in food, and saliva samples were collected for assaying levels of tryptase activity, a possible marker of inflammatory processes. Urine samples were collected prior to the laboratory visit and again at 15 minutes and 50 minutes after challenge. Saliva samples were collected at approximately 50 minutes and 15 minutes pre-challenge and again at approximately 15 minutes and 50 minutes post-challenge. For data analysis, scores for the CPT and the observational ratings were converted to z-scores and aggregated to produce a Hyperactivity Index (HI) score for the pre-challenge and the post-challenge periods [see *Reviewer Comment 11*, below]. The pre-challenge HI was subtracted from the post-challenge HI to produce an HI difference score for statistical analysis. Higher difference scores reflected more negative behavior in response to the Mix B challenge. Mixed Model methods were used to analyze the data [see *Reviewer Comment 12*, below]. Only the results of the behavioral analyses were presented in the final report. The results relating to the metabolic factors were not complete at the time the technical report was prepared and are to be presented at a later date.

3. <u>Modulation of AFCA behavioral effects by genetic polymorphisms</u>

To address the secondary question of whether genetic differences make individual children more or less sensitive to the AFCA (Mix A and Mix B) treatments, selected genetic polymorphisms were analyzed using DNA from cheek cells for all 3/4YO and 8/9YO children of both sexes in the main study [see *Reviewer Comment 13*, below]. Single nucleotide polymorphisms (SNPs) were selected from the dopamine and adrenergic neurotransmitter system since these have previously been implicated in ADHD. Since there is also a suggestion that histamine may be involved in the effects of AFCAs, genetic polymorphisms from this system were also included. Results were presented in the final report for two SNPs in the histamine N-methyltransferase gene (HNMT Thr105lle) and (HNMT T939C), one SNP in the dopamine related catechol-o-methyltransferase gene (COMT Val108Met), and one SNP in the adrenergic neurotransmitter receptor alpha 2A gene (ADRA2A C1292G). The Mixed Models analyses used to determine whether these genotypes modulated the effects of the AFCAs (Mixes A and B) were limited to those children consuming an adequate amount of the challenge *(≥85% consumption* sub-group), since the aim of this analysis was not to establish the impact of the individual additives per se (where the intention to treat based on the *whole sample* is the focus). Thus, the analyses focused on the interactions between genotype and effects of Mix A and Mix B in the *≥85%*

consumption sub-sample of 3/4YO and 8/9YO children.

C. Reviewer Comments Regarding Study Design/Methods

1. ***Questionable accuracy of teacher pre-study behavioral ratings recalling behavior over a period of 6 months***
   The teachers completed behavioral questionnaires prior to the start of the study to indicate frequency of certain behaviors for individual children over the previous 6 months. The accuracy of ratings based on a recall of behaviors over a period of 6 months is highly questionable, particularly for teachers rating each of multiple children in a classroom or playroom setting.

2. ***Questionable use and accuracy of pre-study behavioral ratings by participating parents***
   As noted above for the teacher screening data, the accuracy of pre-study ratings based on a recall of behaviors over a past 6 month period would be questionable. Also, the information derived from these parental ratings did not appear to be used for any analyses in the study itself.

3. ***Disproportionate increases were made in levels of artificial colors in the mixes (A and B) used for the 8/9YO children relative to the mixes (A and B) used for the 3/4YO children, and no adjustments were made to levels of sodium benzoate in either Mix A or B for either age group of children***
   As stated in the technical final report (p. 10), the levels of colors in Mix A used for the 8/9YO children were increased by 1.25 times the levels of colors in Mix A used for the 3/4YO children. However, Mix B levels of colors for the 8/9YO children were disproportionately increased by 2.08 times the levels in Mix B used for the 3/4YO children. Although the rationale for adjusting levels of colors in the mixes for both age groups of children included attempts either to reflect the current average daily consumption of food additives by 3-year-olds in the UK (Mix B) or to account for the increased amount of food typically consumed by children in the 8 to 9 year old range, there was no explanation why comparable adjustments in sodium benzoate levels were not made across mixes or across age groups. The differences in levels and types and quantities of colors and the lack of adjusted levels of sodium benzoate across mixes (A and B) and particularly across age groups essentially precludes any reliable dose response evaluations or comparative assessments of treatment effects between Mix A and Mix B within or between age groups of children.

4. ***Composition of placebo is unclear***
   Other than the placebo drink being made of various proportions of fruit juices, no specific information was provided about the composition of the placebo (in Annex 2, p. 121 of final report one mention was made of two 8/9YO children excluded from the study due to allergic reaction to *blackcurrant juice*). The placebo is a critical component of this challenge study and it would be important to have information about the types of fruit juice used (commercial or fresh). If commercial, what brands were used and did they contain any types of additives? If fresh, how was the juice made and when was it prepared relative to use in the study? Such information about the placebo source and composition are very important in helping to assess the outcome information from this study.

5. ***Parents were not asked about the blinding***
   Although it was determined that the active and placebo juice drinks could not be differentiated based on look and taste and the parents were blind to which challenge drink (active or placebo) was being delivered in any particular week, the investigators did not specifically ask the parents whether they could tell when the active challenge was being given at any time during the study.

6. ***Timing of juice consumption at home***
   While all children in both age groups consumed the juice drinks at home, they consumed the drinks either prior to or after their session at EYS or school (8/9YO children consumed their juice mainly after returning from school). However, there is no information provided as to the percent of children that consumed the juice drinks before going to EYS or school or going for classroom observation or CPT session. This information would be important for interpreting detection of treatment effects relative to latency to onset of effects after dosing.

7. ***Precision of recall data in conduct of weekly parent/teacher ratings and availability of standard ratings***
   The parent and teacher weekly ratings during the study proper were based on recall of behaviors for individual children over the previous week. Recalling behaviors over a previous seven day period, although acceptable, will invariably be less precise than daily ratings. However, daily ratings would have presented logistical difficulties. The investigators did not provide comparative standard teacher/parent behavioral ratings for ADHD and non-ADHD children, with which to gauge the significance of the ratings for the children in this study.

8. ***Reliability of modified Classroom Observation Code (COC) used in 3/4YO children is unclear***
   On page 131 of the final report, the COC used for the 8/9YO children was noted as having adequate inter-observer reliability and ability to discriminate between hyperactive and non-hyperactive children (no definition or criteria provided for designation of ADHD). Since a modified version of the COC was used for the 3/4YO children, it is not clear whether this modified COC was equally reliable. The investigators did not provide comparative standard COC ratings for ADHD and non-ADHD children, with which to gauge the significance of the COC for the children in this study.

9. ***Questionable adequacy of duration used for classroom observation***
   Each child's weekly classroom observation score was based on only three 8-minute observation periods (total of 24 minutes of observation per week). It is questionable whether observation for 8 minutes, three times a week is an adequate sampling frequency or observation duration to provide a reliable measure of representative classroom behavior over an entire week.

10. ***Model 1 is not useable for analysis of data***
    Since statistical Model 1 does not adjust for the potential confounds in the study, the analyses using Model 1 do not provide reliable information about treatment related effects and are of no apparent value in the final evaluation of findings.

11. ***In the "acute challenge phase" of this study the disaggregated behavior measures comprising the Hyperactivity Index (HI) were not analyzed separately***
The component behavior measures, CPT and independent behavior ratings were aggregated to form the Hyperactivity Index (HI). However, these component behavior measures were not analyzed separately (disaggregated) to determine whether changes in CPT performance or the behavior rating contributed more, less or equivalently to treatment related changes in the overall aggregate HI scores.

12. ***The Mixed Model analyses of the "acute challenge" data did not appear to control for potential confounding factors***
In describing the acute challenge findings, the final technical report does not indicate that any potential confounding factors were either identified or controlled for in the Mixed Model methods used to analyze the data. In the absence of controlling for confounding variables there is low confidence in the reliability of the statistical analyses of the acute challenge study findings.

13. ***No description of materials or methods was provided for the genotyping analyses***
The investigators did not provide any details in the final report about the manner in which DNA samples were taken, at what period during the main study samples were taken, how the samples were maintained, or how the samples were analyzed for SNPs. Without such methodological information it was difficult to assess the adequacy of the procedures used in this analysis or the reliability of the data collected.

IV. Study Investigators' Reported Findings

A. Reported Study Findings Relative To Primary Research Question: Do Mixtures of Certain Artificial Food Colors and Sodium Benzoate Increase the Mean Level of Hyperactive Behavior (*Hyperactivity Profile Behaviors*) In Children From the General Population?

1. Study Sample and Background Characteristics
Table 2 shows the numbers of 3/4YO and 8/9YO subjects in the *whole sample*, in a sub-group of children who had *consumed ≥85%* juice drinks in any challenge week over the period of the study, and in a smaller sub-group of children who consumed ≥85% juice drinks and also had no missing behavior data (*complete case* group).

TABLE 2. Numbers of 3/4-Year Old and 8/9-Year Old Children in Study

	Children in Whole Sample N	Children Consuming ≥85% of Challenge Drinks N*	Children in Complete Case (≥85% Consumption & No Missing Data) N*
3/4-Year Olds	153 (79 m/ 74 f)	133	73
8/9-Year Olds	144 (75 m/ 69 f)	119	91

* The sex distributions in the *≥85% Consumption* group and the *Complete Case* group were not specified.

The analysis of background characteristics, including race, marital status, social class/employment of father and of mother, and mother's education level, revealed no significant differences between these three groups or between the groups of children assigned to receive the challenge drinks in different orders over the course of the study.

Based on pre-study teacher's behavioral ratings (ADHD Rating Scale-IV/Teacher version), the sample populations of male and female subjects in both the 3/4YO and 8/9YO age groups were representative of the general populations of age-matched children within the early year settings and schools in terms of the Teacher scores for inattention, hyperactive and total behaviors. As expected, significant gender differences were found in each age group for both the population ($p<0.001$) and study sample ($p<0.001$) with boys having higher behavioral scores than girls.

2. Juice consumption and dietary infractions

Juice consumption over the period of the study remained at what was considered an acceptable level for the majority of children in both age groups. Of the children who completed the study 93% of the 3/4YO children and 85% of the 8/9YO children consumed more than two thirds of all drinks and 75% of both age groups consumed ≥85% (at least 6 out of 7 daily drinks per week). Dietary infractions during the study were considered acceptably low for both age groups with approximately 30% having 0 infractions and 17% having more than 4 infractions. Rates were comparable during active and placebo weeks.

3. Effects of Challenge

The unadjusted mean GHA scores for 3/4YO and 8/9YO children by challenge type are presented in Table 3.

TABLE 3. Unadjusted Mean GHA Scores* for 3/4YO and 8/9YO Children by Challenge Type

		Mix A n	Mix A Mean (SD)	Mix B n	Mix B Mean (SD)	Placebo n	Placebo Mean (SD)
3/4 Year Old Children	Whole Sample	131	- 0.11 (1.03)	134	- 0.14 (1.03)	129	- 0.32 (1.11)
	≥ 85% Consumption	104	- 0.11 (1.03)	108	- 0.15 (1.07)	99	- 0.39 (1.07)
	Complete Case	73	- 0.14 (1.04)	73	- 0.26 (1.05)	73	- 0.44 (0.98)
8/9 Year Old Children	Whole Sample	132	0.25 (0.97)	133	0.33 (1.10)	127	0.19 (1.03)
	≥ 85% Consumption	104	0.26 (0.93)	112	0.32 (1.09)	103	0.19 (1.04)
	Complete Case	91	0.27 (0.92)	91	0.35 (1.08)	91	0.19 (1.06)

* A numerically higher GHA score indicates a greater level of hyperactivity (*hyperactivity profile behaviors*)

It should be noted that the GHA scores (which indicate level of behavior relative to T0 baseline prior to removal of AFCA from the diet and prior to any challenge drinks being administered) for the 3/4YO children in all groups were generally below the baseline level (i.e., less *hyperactive profile behaviors* relative to the T0 baseline). This was interpreted to mean that the effects of reduced hyperactive profile behaviors from withdrawal of AFCA from the diet were not being counteracted by the effects of subsequent AFCA challenges. For the 8/9YO children, however, the GHA scores overall tended to increase above the baseline level (i.e., more *hyperactive profile*

behaviors relative to baseline). This was attributed to an overall decreasing level of CPT performance across the study period for this age group of children, which contributed to an overall increase in their GHA scores (CPT was not administered to the 3/4YO children).

In identifying potential moderating or confounding effects of variables on the behavioral response to challenge, preliminary Mixed Models analyses showed that for both age groups of children the baseline (T0) GHA behavior score was related to the challenge GHA at all subsequent time points and that gender and pre-study diet (3/4YO children only) were individually related to baseline GHA but in no case was there a significant interaction between these or other factors and challenge type (i.e., AFCA or placebo). There was no effect of carryover from challenges in previous weeks on behavior during subsequent challenge weeks. While the 3/4YO children showed no effect of time (week) on the GHA score, for the 8/9YO children there was an effect with GHA increasing across weeks during the study. An examination of the component behavior measures of the GHA indicated that this was due to a gradual worsening of the children's scores on the CPT over the weeks of the study. This was attributed to the children becoming less motivated to perform this intrinsically boring task with repeated testing. The declining CPT performance week by week resulted in increasing GHA scores for the 8/9YO children [see *Reviewer Comment 1*, below].

The final assessment of significance of the effects of challenge (Mix A or Mix B versus placebo) on GHA is based on a Mixed Model analysis with potential confounding factors controlled, including the effects of week during study, gender, GHA in baseline week (T0), number of additives in pre-test diet, maternal education level and social class. The results from the analyses of estimate effect sizes (Mix A versus placebo and Mix B versus placebo) for both 3/4YO and 8/9YO groups of children are shown in Table 4, separately for the full sample (primary outcome), and *post hoc* analyses of two sub-samples one with a high *≥85% consumption* of juice drinks in any challenge week and the other, the *complete case* sub-sample, with ≥85% consumption in all challenge weeks and no missing behavioral GHA scores [see *Reviewer comment 2,* below].

TABLE 4. General Linear Mixed Models estimates of GHA effect size for challenge Mix A versus Placebo and challenge Mix B versus Placebo
[Whole Sample, *>85% Consumption*, and *Complete Case* Groups of 3/4 Year Old and 8/9 Year Old Children with all potential confounding factors controlled]

Subjects	Challenge Type	Estimate Effect Size (95% CI)					
		N	Whole Sample	N	≥85% Consumption	N	Complete Case
3/4 Year Old Children	Mix A vs Placebo	140	0.20 (0.01 to 0.39)*	130	0.28 (0.05 to 0.51)*	73	0.32 (0.05 to 0.60)*
	Mix B vs Placebo	140	0.17 (- 0.03 to 0.36)	130	0.19 (-0.04 to 0.41)	73	0.21 (- 0.06 to 0.48)
8/9 Year Old Children	Mix A vs Placebo	136	0.08 (- 0.02 to 0.17)	119	0.09 (- 0.01 to 0.19)	91	0.12 (0.02 to 0.23)*
	Mix B vs Placebo	136	0.12 (0.03 to 0.22)*	119	0.15 (0.05 to 0.25)**	91	0.17 (0.07 to 0.28)**

* $p<0.05$, ** $p<0.01$

All significant treatment related effects on behavior relative to placebo in both age groups of children (Table 4) were in the direction of increased hyperactivity profile behaviors. For the 3/4YO children, the primary analysis of the data using the *whole sample* showed a statistically significant effect of Mix A on GHA compared to placebo ($p<0.05$) but no significant effects of the Mix B challenge. When the analysis was limited either to those children with *≥85% juice consumption* or to the *complete case* sub-samples, the adverse effect of Mix A on behavior remained statistically significant ($p< 0.05$) and Mix B still had no significant effects.

For the 8/9YO children, the primary analysis of the *whole sample* showed a statistically significant effect of Mix B on GHA compared to placebo ($p<0.05$) but not Mix A. The same pattern of effect occurred when the analysis used the data from the *≥85% consumption* sub-group, that is the effect of Mix B remained significant ($p<0.01$) but Mix A had no significant effects. When the analysis was limited to the *complete case* sub-sample with ≥85% consumption and no missing GHA data, the effects of both Mix A and Mix B were statistically significant ($p<0.05$ and $p<0.01$, respectively) [see *Reviewer comment 3,* below].

B. Reviewer Comments Regarding Reported Study Findings for Primary Research Question

1. *Limited information regarding CPT performance decrements of 8/9YO children over study period*

CPT performance of the 8/9 year old children declined over the course of the study. Although little detailed information was presented, the study investigators concluded that this decline was attributable to the children becoming bored with the repeated conduct of this inherently tedious task. Sufficient additional information should have been provided to support this conclusion.

2. *Problematic aspects of the data analyses*

The analyses in this study were replicated in each age group for the *full sample* of subjects, for a high consumption group *(≥85% consumption* of juice drinks in any challenge week), and for a *complete case group* (*≥85% consumption* in all challenge weeks and no missing GHAs). The rationale for including the latter two groups was to determine whether non-compliance (failure to consume challenge drinks) and the method of handling missing data affected the pattern of results. As shown in the findings from the analyses of the overall GHA scores, there was some variability in the occurrence of statistically significant effects across subgroups. Specifically, when the *whole sample* or *≥85% consumption* groups of 8/9YO children were used for analysis, only Mix B had a significant effect. However, when the *complete case* group of 8/9YO children was used, both Mix A and Mix B had significant effects on GHA scores. Additional more notable examples of variability across subgroups occurred in the analyses of the change in GHA scores for the component (disaggregated) behavior measures (shown in Section IV, C below). Yet, there was little, if any, qualification accompanying the description of the data analyses or the results using the whole sample or the two sub-samples of subjects. This resulted in confusion as to which analyses provided the primary study results. However, the principle investigators eventually stated in the Discussion of the Primary Research Question Findings (p. 26) that the use of the "*whole sample*" of the 3/4YO and of the 8/9YO children is considered to be "the primary analysis of the data on an intention

to treat basis". While the variability in significant treatment effects using the different subgroups of subjects for analysis may indicate that non-compliance in consumption of challenge drinks or the method of handling missing data affected the pattern of results, the investigators do not discuss the impact of this problematic aspect of the data analyses on interpretation of the study findings.

3. ***There is no clarification of the behavioral changes in this study as necessarily "adverse" (p.22) or "deleterious" (p.26)***
There was no information provided in this study to suggest that the changes in behavior based on parental reports were adverse, detrimental or maladaptive in any demonstrable manner. The children's behavior under challenge conditions appeared to be within the range of behavioral levels exhibited by the general population of age-matched children. The significant statistical findings should be presented simply as "effects" and not "adverse" or "deleterious" effects, without further clarification.

C. Reported Study Findings Regarding Secondary Research Questions

1. (Secondary Research Question) Effects of the AFCA challenge mixes on component behavior measures of GHA

To gauge the extent to which the individual behavior components of the overall GHA may have contributed to the effects of AFCA treatment on children's behavior, additional statistical analyses were conducted as a secondary outcome measure. These analyses considered the effect of challenge on the disaggregated standardized GHA scores for each of the component behavior measures; teacher ratings, parent ratings, classroom observation scores, and CPT. The results of these analyses are presented in Table 5.

In presenting the results of analysis of the effects of challenge on the disaggregated GHA scores for individual behavior measures, the principle investigators state that any single indicator is likely to be relatively less reliable compared to the overall composite GHA. The consequent increased measurement error in the analysis of individual behavior measures makes it less likely that a significant effect will be detected. For this reason the principle investigators concluded that the results for the disaggregated behavior measures are most appropriately discussed in terms of the effect sizes of Mix A versus placebo and Mix B versus placebo [see *Reviewer Comment 2(a)*, below]. With specific reference to the 3/4 YO children (p.104), the investigators suggested that the findings should be viewed in the context of the additive mixes being consumed at home and not at the early years setting (or in the classroom observation setting) and that any behavior score will be a function, among other things, not only of the amount of juice consumed but also the time and individual differences in absorption of additives.

TABLE 5. Effect of sizes of Mix A versus Placebo and Mix B versus Placebo for each disaggregated behavior measure
[*Whole Sample*, *≥85% Consumption*, and *Complete Case* Groups of 3/4 Year Old and 8/9 Year Old Children with all potential confounds controlled]

		Mix A vs Placebo (Estimate Effect Size)			Mix B vs Placebo (Estimate Effect Size)		
		Whole Sample	**≥85% Consumption**	**Complete Case**	**Whole Sample**	**≥85% Consumption**	**Complete Case**
3/4 Year Olds	Parent rating	0.33	0.49 (p<0.016)	0.55 (p<0.027)	0.27	0.36	0.37
	Teacher rating	0.01	0.03	0.09	0.06	0.08	0.10
	Classroom Observation	0.09	0.10	0.08	0.001	- 0.01	- 0.02
8/9 Year Olds	Parent rating	0.01	0.03	0.03	0.13 (p=0.031)	0.13 (p=0.046)	0.08
	Teacher rating	- 0.04	- 0.01	0.00	- 0.03	0.01	0.04
	Classroom Observation	0.02	0.08	0.04	0.01	0.05	0.07
	Continuous Performance Task (CPT)	0.10	0.08	0.18	0.19	0.20	0.31 (p=0.015)

As presented in Table 5, the analyses of the challenge versus placebo effect sizes for the individual component measures of behavior were analyzed for *the whole sample*, for the subset *consuming ≥85%* or more of the drinks, and for the *complete case* sub-sample of 3/4YO and 8/9YO children [see *Reviewer Comment 2(b),* below]. Overall, the largest effects in the 3/4YO children were found for the parental ratings with both Mix A and B challenges. For the 8/9YO children the largest effects overall were found for the computerized test of attention (CPT) with both Mix A and B challenges, and moderate effects were found for parental ratings with Mix B challenge. In general, for both age groups a majority of the challenge versus placebo effect sizes for all behavior measures under both Mix A and Mix B challenge conditions were in the direction of increased hyperactivity (*hyperactivity profile behaviors*).

Specifically, for the *whole sample* of 3/4YO children the largest effects were those based on parental reports with both Mix A and Mix B challenges, neither of which was statistically significant. For the *whole sample* of 8/9YO children the largest effects were found for the computerized test of attention (CPT) under both Mix A and Mix B challenges, neither of which was statistically significant, but the moderate effects based on parental ratings were significant (p<0.04) for Mix B but not Mix A. Notably, negligible non-significant effect levels were found for teacher ratings and classroom observations in the *whole sample* of both 3/4YO and 8/9YO children.

For the *≥85% consumption* and the *complete case* groups of 3/4YO children the largest effects were based on parental reports for both Mix A and Mix B, with only the parental report effect size for Mix A being statistically significant in both the *≥85% consumption group* (p<0.02) and the *complete case* group (p<0.03). For the 8/9YO children the largest effects were still found in the CPT under both Mix A and

Mix B challenges, with only the CPT effect size for Mix B in the *complete case* group being statistically significant ($p<0.02$). In the analyses of the *≥85% consumption* and the *complete case* groups of 3/4YO and 8/9YO children, the only other statistically significant effect was the parental report effect for Mix B in the *≥85% consumption* group of 8/9YO children ($p<0.05$). Negligible non-significant effect levels were still found for teacher ratings and classroom observation ratings using either the *≥85% consumption* or the *complete case* sub-samples of 3/4YO and 8/9YO children.

2. *Reviewer Comments:* Effects of the AFCA challenge mixes on component behavior measures of GHA

(a) *Complementary changes across disaggregated behavior measures can provide reliable confirmation of significant treatment effects.*

The principle investigators minimized the significance and utility of disaggregated measures by stating that they consider analysis of disaggregated measures to be less sensitive for finding a statistically significant effect and thereby less reliable than the global aggregate measure, GHA. They state that "Any single indicator is likely to be relatively less reliable compared to the aggregate measure. The consequent increased measurement error makes it less likely that a significant effect will be detected." However, it should be noted that analyses of disaggregated behavior measures were sufficiently sensitive to provide a number of statistically significant findings. The statistical argument proposed by the principle investigators appears to support the use of artificially enhanced detection sensitivity by combining different types of behavioral measures (parent ratings, teacher ratings, classroom observation codes, and continuous performance testing), which effectively serves to lower the error term and enhance the likely occurrence of statistical significance. However, sensitivity (whether artificially increased or not) is not the only, nor is it arguably the most important, aspect of reliability in detecting a true treatment related effect. Replication of findings and concordance of changes across the various behavior measures, which serve to help confirm the occurrence of treatment effects, are also important elements of reliability. The present study focuses on "hyperactivity" as the target behavior. As used by the investigators in this study, the term "hyperactivity" indicates a behavioral profile characterized by overactive, impulsive and inattentive behavior. Finding complimentary statistically significant treatment related changes in the hyperactivity profile across several of the various behavior measures would seem to be more informative about the scope of effect and provide more confidence of confirmation of relevant treatment effects than a statistical finding with the enhanced sensitivity of a single aggregate score. No such complimentary treatment effects were found.

(b) *Use of the three sample groups of children for analysis of the component (disaggregated) behavior measures results in variable statistical outcomes but relevance of this to data interpretation are not discussed*

[See also *Reviewer Comment in section IV, B, 2* above regarding analysis of overall GHAs in main study]. As seen in the analyses of the treatment-related changes in the disaggregated GHAs for the component behavior measures, there is notable variability in the outcome of statistically significant effects depending upon whether the *whole sample,* the *≥85% consumption* sub-sample, or the *complete case* sub-sample of children was used for the analysis. For example, based on analysis of the component behaviors using the *whole sample* of children, there were no significant effects for the 3/4YO children, and for the 8/9YO children only the parent ratings were significantly affected with Mix B challenge. None of the teacher ratings or

classroom observation ratings (both age groups) or CPT scores (8/9YO) were significantly affected. However, when the analyses used the *complete case* group of children, parent ratings were significantly affected only for the 3/4YO children with Mix A, and CPT scores were significantly affected but only for the 8/9YO children with Mix B. The teacher and classroom observation ratings were still not significantly affected. Although this obvious variability in significant treatment effects across subgroups of subjects may indicate that non-compliance in consumption of challenge drinks or the method of handling missing data affected the pattern of results, the investigators do not discuss the impact of this problematic aspect of the data analyses on the confidence in the statistical significance of the study findings and their interpretation.

3. **(Secondary Research Question) <u>Behavioral effects and metabolic mediators following acute challenge with AFCA – Study Findings and Interpretation</u>**
The acute challenge component of the project explored the possibility of demonstrating short term changes in hyperactive behavior (*hyperactivity profile*) immediately post-challenge with Mix B for two groups of 8/9YO boys who did (n=15) or did not (n=15) respond to Mix B in the main challenge study. The results of the Mixed Models analyses with groups combined showed no statistically significant main effect of challenge with Mix B on the Hyperactivity Index (a composite aggregate of the CPT scores and the observational ratings) compared to Placebo (p=0.096), no significant difference between the Hyperactivity Index levels of responder and non-responder groups (p=0.134), and no significant Challenge x Group interaction on the response to Mix B challenge (p=0.072). Although all elements of the analyses were non-significant, there was a general trend towards increased *hyperactive profile* behavior in response to Mix B challenge, particularly by the group of "responders" compared to the "non-responders" [see *Reviewer Comment (a)*, below].

The principle investigators contended that the findings from the acute challenge study suggest that the 8/9 year old "responders" to challenge with Mix B exhibit effects within a short period of time after dosing (approximately one hour). The children in the main study consumed their challenge drinks at home. The 3/4 YO children consumed their drinks either prior to or upon returning from the early year setting but the 8/9 YO children usually consumed their drinks upon return from school (percentages were not reported). With an approximate one hour latency to onset of acute effects, the investigators suggested that this makes it likely that the treatment-induced behavior changes occurred in the home setting and this may be an explanation as to why the strongest effects were found for the parental ratings but not for the school or early years settings based measures [see *Reviewer Comment (b),* below].

4. ***<u>Reviewer Comments</u>*<u>: Behavioral effects and metabolic mediators following acute challenge with AFCA – Study Findings</u>**

(a) ***<u>Statistical analyses of data in the "acute challenge" component study did not control for confounding variables</u>*.**
None of the confounding variables which were controlled in the data analyses of the primary study, such as CPT aggregate in baseline week, number of additives in pre-trial diet, maternal educational level and social class, appear to have been

incorporated into the Mixed Model methods used for analysis of the acute challenge data. Since potential confounds were not controlled in the data analysis, this calls into question the reliability of the findings from the acute challenge component study.

(b) ***Rationale for explaining why treatment effects in the main study were detected only in parental ratings was based on inaccurate interpretation of acute study findings.*** Aside from the questionable reliability of the data analysis in the acute challenge study and the fact that no significant treatment effects were found, the trends in the acute challenge study (although not statistically significant) suggest that the onset of response following ingestion of Mix B challenge occurred within a short period of time (an hour). Since all of the children in the primary study consumed their challenge drinks at home (the 8/9 YO children usually consumed their challenge drinks after returning from school and the 3/4 YO children consumed their drinks either prior to or upon returning from their school setting), it seems likely that the parents in the home setting would have observed behavioral changes that might have occurred in the children. However, the acute challenge study was not designed to determine the duration or latency of treatment effects after challenge.

5. **(Secondary Research Question) Modulation of AFCA behavioral effects by genetic polymorphisms – Study Findings**
The results of the analyses examining the modulation of the effect of challenge by the children's genotype are shown in Table 6 for the 3/4YO children and Table 7 for the 8/9YO children. The analysis of the datasets for each genotype included an initial test for the main effect of challenge (Mix A and Mix B) on GHA levels, independent of whether the genotype was present or absent. In the 3/4YO children there was a significant main effect of challenge on GHA levels with both Mix A and Mix B in the analyses for HNMT Thr105lle (p=0.004 and p=0.02, respectively) and HNMT T939C (p=0.005 and p=0.036, respectively), but there was no significant main effect with either challenge mix in the analysis for COMT Val108Met or ADRA2A C1291G. Comparably, in the 8/9YO children the main effect of challenge on GHA levels was significant with both Mix A and Mix B in the analysis for HNMT Thr105lle (p=0.046 and p=0.001, respectively) and with Mix B (but not Mix A) in the analysis for ADRA2A C1291G (p=0.036). There was no main effect of challenge with either Mix A or Mix B in the analysis for the COMT Val108Met dataset. Although these results for main effect of challenge were not discussed to any great extent by the principle investigators in the final report, it seemed appropriate for purposes of completeness to describe these particular results in more detail in this review [see *Reviewer Comment (a),* below].

TABLE 6. Mixed Models analyses of interactions between Genotype and Behavioral Responses (GHA) to Additive Mix A and Mix B versus Placebo in the ≥85% consumption sub-sample of 3/4 YO Children with All Potential Confounds Controlled

		[Effect Size (p value)]			
		HNMT Thr105Ile	**HNMT T939C**	**COMT Val108Met**	**ADRA2A C1291G**
Additive Challenge	**Mix A vs Placebo**	0.39 (*p=0.004)	0.42 (*p=0.005)	0.39 (ns: p=0.087)	0.27 (ns: p=0.110)
	Mix B vs Placebo	0.30 (*p=0.020)	0.30 (*p=0.036)	0.11 (ns: p=0.645)	0.10 (ns: p=0.540)
	Summary	***Both Mix A and Mix B increase GHA levels***	***Both Mix A and Mix B increase GHA levels***	*GHA levels not affected by either Mix A or B*	*GHA levels not affected by either Mix A or B*
Genotype	**allele present vs allele absent**	0.51 (*p=0.021)	0.38 (ns: p=0.071)	-0.17 (ns: p=0.458)	-0.24 (ns: p=0.222)
	Summary	***GHA levels are higher with allele present***	*GHA levels and allele presence/absence not related*	*GHA levels and allele presence/absence not related*	*GHA levels and allele presence/absence not related*
Challenge (Mix A or Mix B) x Genotype (allele present or absent):	**Mix A vs Placebo**	-0.53 (*p=0.041)	-0.46 (ns: p=0.061)	-0.23 (ns: p=382)	0.01 (ns: p=0.959)
	Mix B vs Placebo	-0.40 (ns: p=0.134)	-0.23 (ns: p=0.338)	0.12 (ns: p=0.662)	0.20 (ns: p=0.389)
	Summary	***Absence of allele enhances GHA response to Mix A but not to Mix B***	*GHA response to Mix A and B not affected by allele*	*GHA response to Mix A and B not affected by allele*	*GHA response to Mix A and B not affected by allele*

TABLE 7. Mixed Models analyses of interactions between Genotype and Behavioral Responses (GHA) to Additive Mix A and Mix B versus Placebo in the ≥85% consumption sub-sample of 8/9 YO Children with All Potential Confounds Controlled

		[Effect Size (p value)]			
		HNMT Thr105lle	**HNMT T939C**	**COMT Val108Met**	**ADRA2A C1291G**
Additive Challenge	**Mix A vs Placebo**	0.11 (*p=0.046)	0.19 (*p=0.003)	0.08 (ns: p=0.379)	0.11 (ns: p=0.104)
	Mix B vs Placebo	0.19 (*p=0.001)	0.25 (*p<0.001)	0.14 (ns: p=0.151)	0.14 (*p=0.036)
	Summary	***Both Mix A and Mix B increase GHA levels***	***Both Mix A and Mix B increase GHA levels***	*GHA levels not affected by either Mix A or B*	***Only Mix B increases GHA levels***
Genotype	**allele present vs allele absent**	0.01 (ns: p=0.956)	0.18 (ns: p=0.089)	0.12 (ns: p=0.295)	0.05 (ns: p=0.649)
	Summary	*GHA levels and allele presence/absence not related*	*GHA levels and allele presence/absence not related*	*GHA levels and allele presence/absence not related*	*GHA levels and allele presence/absence not related*
Challenge (Mix A or Mix B) x Genotype (allele present or absent):	**Mix A vs Placebo**	-0.10 (ns: p=0.403)	-0.24 (*p=0.021)	0.02 (ns: p=0.874)	-0.05 (ns: p=0.607)
	Mix B vs Placebo	-0.24 (*p=0.050)	-0.23 (*p=0.026)	0.02 (ns: p=0.865)	-0.004 (ns: p=0.967)
	Summary	***Absence of allele enhances GHA response to Mix B but not to Mix A***	***Absence of allele enhances GHA response to both Mix A and Mix B***	*GHA response to Mix A and B not affected by allele*	*GHA response to Mix A and B not affected by allele*

There were no significant main effects of any genotype on GHA at baseline (T0) in either age group of children (data not shown). During the challenge study, only one genotype, HNMT Thr105lle, in the 3/4YO children showed a significant (p=0.02) main effect on GHA levels, which were higher with Thr105lle/*present* compared to Thr105lle/*absent*. The relevance of this is not known.

The results of the interaction analyses showed that the COMT Val108Met and ADRA2A c1291g polymorphisms apparently had no modulating influence on the effects of AFCAs (Mix A and B) on GHA levels in either the 3/4YO or 8/9YO children. However, significant modulating effects were found for both HNMT Thr105lle and the HNMT T939C polymorphisms in both the 3/4YO and 8/9YO children.

Specifically, for the 3/4YO children a modulating effect of the HNMT Thr105lle/*present* genotype was found which significantly (p=0.04) reduced the adverse effects of Mix A, that is, fewer *hyperactivity profile behaviors* were seen in response to Mix A challenge [see *Reviewer Comment (b),* below]. A similar, but nonsignificant (p=0.06), moderating effect of the HNMT T939C/present genotype on

the adverse effects of Mix A was noted. Neither the HNMT Thr105lle nor the HNMT T939C genotype significantly influenced the effects of Mix B in the 3/4YO children.

For the 8/9YO children the HNMT Thr105lle/*present* genotype significantly (p=0.05) reduced the adverse behavioral effects of Mix B, but not Mix A. The HNMT T939C/*present* genotype had an even greater modulating influence, significantly reducing the adverse effects of both Mix A (p=0.02) and Mix B (p=0.026).

In the discussion of these genotype findings the principle investigators suggest a link between histamine and hyperactivity with certain polymorphisms in the HNMT gene moderating behavioral responses to the mixture of food colorings and the benzoate preservative present in Mix A in the 3/4YO children and in Mix A and B in the 8/9YO children. They also suggest that the current focus on dopamine in studies of ADHD needs to be extended to histamine. HNMT polymorphisms impair histamine clearance. The presence of H3 receptors in the brain provides a possible mechanistic explanation for the interactive effects found. Many environmental factors can increase histamine release, including infections as well as many food items and certain artificial food colors. The authors indicate that this would explain the frequent claim that food allergy/intolerance is a cause of hyperactivity and the effects of infections in aggravating aberrant behavior. This clearly indicates a potential target for therapeutic intervention in ADHD focused on the H3 receptor [see *Reviewer Comment (c),* below].

6. *Reviewer Comments*: Modulation of AFCA behavioral effects by genetic polymorphisms – Study Findings

(a) *Variability in finding of significance for main effect of challenge (Mix A and Mix B) on behavioral responses (GHA) with repeated Mixed Models analyses for different genotypes raises questions about replicability of the Mixed Models analyses and uncertainties about the findings of significant treatment effects in the primary study.*

For each age group, the *≥85% consumption* sub-sample of subjects identified in the primary study was used for determining the modulating influence of four different genotypes on the behavioral effects of treatment (Mix A and Mix B). Four separate datasets each comprised of virtually all of the same subjects were used for the four genotype analyses. Each of the four datasets used for the four genotype analyses was initially tested (Mixed Models analysis) for the main effect of challenge on GHA levels, independent of whether the genotype allele was present or absent, and subsequently tested for Challenge x Genotype (polymorphism *present* or *absent*) interactions. Effectively, the same Mixed Models analysis for a main challenge effect was repeated four times using separate datasets comprised of basically the same population of experimentally treated subjects; thus, comparable statistical findings with regard to main treatment effects (independent of genotype being present or absent) should occur across all four genotype analyses. In fact, however, the statistical findings regarding main challenge effects were notably inconsistent across the four genotype analyses. In the 3/4YO children significant challenge effects were found for Mix A and Mix B in the analyses of the two HNMT Thr105lle and HNMT T939C datasets, but not in the analyses (using essentially the same sample of children) of the COMT Val108Met and ADRA2A C1291G datasets. Comparably, in

the 8/9YO children significant challenge effects were found for Mix A and Mix B in the analyses of the HNMT Thr105lle and HNMT T939C datasets and for Mix B in the analyses of the ADRA2A C1291G dataset, but not in the analyses of the COMT Val108Met dataset. Even though these analyses were conducted using only the *≥85% consumption* sub-sample of children, such variability in the statistical findings raises questions regarding the replicability of the Mixed Models method at least as used in the determination of the modulating influence of genotype on the behavioral effects of AFCA. A specific evaluation by a statistician of the adequacy and replicability of the statistical procedures used would be appropriate.

(b) ***The moderating influence specifically of theThr105lle/present genotype decreasing the behavioral effects of Mix A in 3/4 YO children appears contradictory to the main effect of genotype on GHA levels.***

The association of Thr105lle/*present* with a reduction in the behavioral effects of Mix A in 3/4 YO children appears contradictory, since a main effect analysis showed this same genotype, Thr105lle/*present,* to be significantly associated with higher overall GHA levels compared with Thr105lle/*absent* genotype. The investigators attempted to address this by suggesting that the two histamine risk alleles in this study have two actions. The first is to influence the overall level of GHA, significantly for the younger children, and second to make the children more vulnerable to the effects of AFCAs on behavior. They add that the role of genes in influencing behavior needs to be understood not by just their main effects of raising levels, for example, of hyperactivity (*hyperactivity profile behaviors*) but also by the interplay both with each other in gene-gene interactions and also by interactions with environmental factors such as diet. The above suggested effects for the histamine risk alleles are not completely consistent with the finding that there were no main effects for either histamine polymorphism on GHA levels in the 8/9 YO children, yet both Thr105lle and T939C were shown to have had significant modulating effects on the behavioral responses to AFCA challenges.

(c) ***The findings from the analysis of the genotype component study data which involved children from the general population were inappropriately extrapolated to ADHD children.***

The suggested utility of genotype results from this study of behavior in a general population of children to studies dealing with the specific condition of ADHD is an inappropriate extrapolation. The present study attempts to provide some information related to whether any effects of AFCAs on behavior are modulated by genetic differences between children in a normal population. This is not the same as a population of ADHD children. ADHD is a specific neurologic disorder characterized by dysfunctional behavior and is not the extreme end of a biological continuum of normal behavior.

7. Additional Secondary Questions – Study Findings

(a) ***Consistency in response between Mix A and Mix B.***

The challenge mixes differed from one another both in terms of the artificial food colors included and in the doses. It is therefore difficult to interpret differential responses by individual children to the two mixes. The distribution of the GHA scores for both mixes at both ages was normal, the effects were on a continuum and there was no immediate evidence in the distribution pattern of a sub-group of children who were distinctively responsive to the mixtures.

(b) ***Dose-response relationship.***
This study was not designed to investigate the effect of dose on response. In fact the variations in dose taken for each mix were not under experimental control but rather arose both from different levels of color additives in each mix and from differential compliance with the consumption of the stipulated dose in the challenge drink. This means that the interpretation of the dose-response relationship is problematic not least in terms of ambiguity in the direction of the effects.

(c) ***Difference between the 3/4 year old and 8/9 year old children in changes in GHA over study period.***
Following the design requirements from the Food Standards Agency, after a baseline period on normal diet, the children were placed on a withdrawal diet without color additives or benzoate preservatives and simultaneously started on a placebo drink. This was done to minimize the placebo effects such that throughout the study the children were receiving a drink of some kind. This meant that the effects of withdrawal were confounded with those of placebo. The pattern of changes in GHA over the period of the study for the 3/4YO children was such that the *hyperactivity profile behaviors* remained below baseline (TO) levels (normal behavior for general population children on regular diets) even though they were exposed to the active challenge mixes for 2 of the 6 weeks of the study period. This was interpreted as the effects of withdrawal which reduced the *hyperactive profile behaviors* such that subsequent challenges to raise levels of these behaviors were counteracted. For the 8/9YO children the GHA tended to increase above baseline level even during periods of placebo challenge. This was deemed due to the inclusion of the Continuous Performance Test (CPT) for the 8/9YO children. This component of the GHA, which was absent for the 3/4YO children, showed progressively worsening scores over the study period, resulting in progressively increasing GHA scores. All other component measures of the 8/9YO children's GHA remained below or close to the baseline values. The principle investigators noted that the CPT time effects were controlled in the Mixed Models analyses reported for this study [see *Reviewer Comment in Section IV,B.1, above*].

V. Study Investigators' Overall Conclusions of Study

The investigators asserted that this study provided evidence that adverse effects of certain mixtures of artificial food colors and benzoate preservative (AFCA) on hyperactivity can be identified in general population samples of 3/4 year old and 8/9 year old children under certain circumstances. They also contended that these findings replicate the adverse behavioral effects of Mix A previously reported on a large sample of 3 year old children from the general population. In their opinion this evidence collectively provides support for the case that certain food additives may exacerbate hyperactive behaviors (*inattention, impulsivity and overactivity*) in some groups of children. The size of the effects of the AFCA on the average hyperactivity score is lower than that reported for clinical samples and the level of individual variation in response was high. The investigators assert that there are major genetic influences on hyperactivity and this study has shown that differential sensitivity to AFCA resulting from selected genetic polymorphisms is one means by which genetic influences on hyperactivity may be mediated. The investigators consider that these findings demonstrate that adverse effects are not found only in those children at the extreme of hyperactivity, namely those diagnosed with ADHD, but can also be found in the general population and across a range of severity of hyperactivity.

VI. Reviewer Conclusions

The primary hypothesis tested in the present study was that mixtures of certain artificial food colors and sodium benzoate increase the mean level of *hyperactivity profile behaviors* of inattention, impulsivity and overactivity in children from the general population. It should be noted that the nature of the "hyperactivity" behaviors and the magnitude of change reported in this study are not associated with or indicative of ADHD, but rather refer to behaviors occurring within the general population. No information was provided in the study to suggest that the behavioral changes noted were adverse, detrimental or maladaptive in any demonstrable manner. Two age groups from the general population were used, a 3/4 YO group and an 8/9 YO group. The study design was a double-blind within-subject cross-over challenge with two active challenge items and a placebo, each administered in randomized order in a fruit juice drink daily for a one week period. The active challenges were two mixtures (Mix A and B), each containing artificial food colors and sodium benzoate. Mix A and B differed from each other in the quantity and composition of the color additives used, but the level of sodium benzoate remained constant. The active mixtures used to challenge the 8/9 YO children, relative to the 3/4 YO children, had appropriately higher quantities of food colors (but not higher levels of sodium benzoate) to account for their larger size. The placebo was a mixture of unspecified fruit juices. A range of weekly behavior measures was used to detect treatment related effects, including parent ratings, teacher ratings, classroom observation ratings, and, for the 8/9 YO children only, a continuous performance test (CPT). However, the primary outcome measure used in this study to assess treatment related changes in *hyperactivity profile behaviors* was a "global aggregated measure of hyperactivity" (GHA) which was derived by combining the standardized z-scores from the different sources of behavior measures. A high GHA indicated greater *hyperactivity profile behaviors*.

Based on the analysis of the *whole sample* of children (considered the primary analysis), the findings from the primary study were that challenge with Mix A of artificial food colors and benzoate preservatives elicited statistically significant increases in GHA levels of 3/4 YO children and that challenge with Mix B elicited statistically significant increases in GHA levels of 8/9 YO children. Additional analyses were conducted to assess the effects of both challenge Mixes on each individual (disaggregated) component behavior measure. The principle investigators attempted to minimize the significance and utility of the individual (disaggregated) behavior measures as being less sensitive and less reliable than the global aggregate measure, GHA. This reviewer does not agree, but is of the opinion that, in order to fully interpret the significance of the treatment-related findings in the primary study, it is important to consider how each of the component behavior measures contributed to the significant treatment effects on the overall GHA levels. Analyses of the disaggregated behavior measures were sufficiently sensitive to identify a number of significant behavioral changes. The finding of complementary significant treatment related changes across several of the various behavior measures would not only provide confirmation of biologically relevant changes but would also provide specific information about the treatment effects that could enhance interpretation of the study findings. The analyses of the individual (disaggregated) component behavior measures for the *whole sample* of subjects indicated that the parental rating was the major contributor for the primary effects of Mix A on GHA levels in 3/4 YO children, although the parental rating measure alone did not show a statistically significant effect of Mix A challenge. In the *whole sample* of 8/9 YO children the parental rating and CPT scores were the major contributors for the primary effects of Mix B on GHA levels but only the parental rating and not the CPT scores showed a statistically

significant treatment effect for Mix B. In neither age group of children were there any notable changes in teacher or classroom observation ratings, indicating that these behavior measures did not confirm the overall treatment effects on GHA levels based on parental ratings. A previous study by Bateman et al (2004), using a large sample of 3 year old children also from the general population, found significant behavioral effects of Mix A based specifically on parental ratings but, similar to the present study, these effects were not confirmed by a clinical behavioral evaluation. Contrary to the contention of the principle investigators, the previous study's significant findings were not completely replicated in the present study, since, as noted above, the analysis of the individual disaggregated component behavior measures revealed that the parental ratings (the primary contributor to the significant effect of Mix A on GHA levels) showed only a non-significant trend for behavioral changes with Mix A challenge.

A rationale was suggested by the principle investigators to explain why treatment effects in the primary study based primarily on the parental ratings were not confirmed by teacher or classroom observation ratings. The principle investigators interpreted trends (findings were not statistically significant) from an "acute challenge" study using 8/9 YO children to suggest a rather short onset time (less than an hour) for the appearance of treatment effects from Mix B challenge. The principle investigators concluded that, since all of the children in the primary study consumed their challenge drinks at home, the short onset time for treatment effects to appear would suggest that it was more likely for the parents in the home setting, rather than teachers or classroom observers, to have seen the behavioral changes in the children. In the opinion of the reviewer, this scenario does not adequately resolve the issue of parental ratings not being confirmed by teacher or classroom observations for two basic reasons. First, the investigators did not consider the importance of when the children consumed their challenge drinks. If the drinks were consumed before going to school the teachers/classroom observer might have been more likely than parents to see the treatment effects. The opposite would be more likely if drinks were consumed after school. Unfortunately, the study report stated only that the 8/9 YO children usually consumed their drinks after returning from school and the 3/4 YO children consumed their drinks either prior to or upon returning from their school setting, but no specific information was given. Second, while the "acute challenge" study suggested a short onset time for the appearance of treatment effects, that study was not designed to determine how long those treatment effects would last. It is conceivable that treatment effects, if any occurred, could have persisted into the day after challenge, when the children returned to the school or early year setting, and if so, should have been detected by the teachers and classroom observers. None of the findings from the "acute challenge" study, therefore, help explain why the overall effects of challenge on GHA levels appeared to be based primarily on the parental ratings for the *whole sample* of both aged groups of children, along with the CPT scores for the 8/9YO children, with only nominal, if any, treatment effects being detected with teacher ratings or classroom observation ratings. Numerous alternative explanations for the lack of confirmation of parental detection of treatment effects by other behavior measures in either age group of children could be speculated. For example, the specific or subtle behavioral changes the parents detected were either not detectable with the other behavior measures or were not expressed in the classroom or other test environments. Consideration should, however, be given to possible alternative explanations, for example that the blind may somehow have been broken, or that the parental findings were simply statistical false positives. In the absence of an adequate explanation, the fact that there was no confirmation of trends for treatment effects between parental ratings and the other behavior measures lessens confidence in the relevance and reliability of the parental based findings.

There are several data analysis issues that may also affect the confidence in the finding of statistically significant overall effects in the primary study including significant effects of Mix A on GHA levels in the 3/4 YO children, and of Mix B in the 8/9 YO children. One issue involves the fact that, in the statistical analyses for treatment related changes in overall GHAs and in the component (disaggregated) behavior measures for both age groups, the finding of statistically significant effects varied notably depending upon whether the *whole sample* or either of the sub-samples *(>85% consumption* group or *complete case* group) was used for the analysis. As noted by the principle investigators, variability in the statistical outcome between the whole sample and sub-sample analyses may indicate that non-compliance in consuming the challenge drinks or the method used to handle missing data may have affected the pattern of results. But, in describing the results the investigators did not discuss or clarify the impact or relevance of these problematic aspects of the data analyses on the confidence in the statistical significance of the study findings and their interpretation. This uncertainty lessens confidence in the reliability and relevance of the statistically significant study findings. A second issue concerns the inconsistency in finding significant main effects of challenge (Mix A and Mix B) on overall behavioral responses (GHA levels) in the analyses of the genotype data (a secondary research component of this study discussed below). In each age group of children a Mixed Models analysis was conducted to test for the influence of each of four genotypes on the behavioral effects of AFCAs. Since the analysis of each of the four genotypes included a test for main effect of challenge (Mix A and Mix B) on GHA levels, the analysis for a main challenge effect was repeated four times using basically the same group of subjects (only the *≥85% consumption* sub-sample was tested). While significant main challenge effects were occasionally detected, these effects were not consistently detected across the four repeated analyses. Although these particular repeated analyses were conducted using the same *≥85% consumption* sub-sample of children, the inconsistency in finding significant main challenge effects raises questions regarding the replicability of the Mixed Models statistical analysis. Since the Mixed Models analysis was used to analyze the primary study data, there are also uncertainties about the replicability and the relevance of the statistically significant challenge effects in the primary study. A specific evaluation by a statistician of the adequacy and replicability of the statistical procedures used for the primary study and the genotype analyses would be appropriate.

The present study also addressed a secondary question involving the possibility that the behavioral effects of AFCAs (Mix A and Mix B) may be moderated by genetic differences among children in a normal population. Four selected genetic polymorphisms were analyzed using cheek cell DNA from all children of both age groups. The genetic data from the *>85% consumption* sub-sample of children were evaluated for interactive effects on the AFCA related changes in behavior. Suggestive findings were presented that supported some level of genetic influence (specifically two histamine N-methyltransferase gene polymorphisms, Thr105lle and T939C) on the sensitivity to the behavioral effects of certain AFCAs in children from the general population. Due to the fact that no methodological details were provided for the genotyping analyses and the questions regarding the statistical analysis used for these data, it was difficult to assess the adequacy of the procedures used in this analysis or the reliability of the findings. These interesting but preliminary genotype findings should be replicated and extended in a more focused and well-designed study with more detailed information about the procedures used.

Overall, the primary study findings are suggestive of low level behavioral effects of AFCAs on behavior in 3/4 YO children (Mix A) and 8/9 YO children (Mix B), limited to detection based primarily on parental ratings and possibly CPT scores for the older children. However, due to the absence of confirmation of treatment effects between parental ratings and other sensitive

behavior measures together with the concerns about the data analyses described above and various procedural weaknesses (outlined in the section on Study Weaknesses, below), it is the opinion of this reviewer that there is questionable confidence in the reliability and biological relevance of the primary findings from this study.

VII. REVIEWER NOTED STUDY STRENGTHS

(1) The recruitment procedure enabled the selection of two study groups of children, one comprised of 3/4 YO and the other 8/9 YO, drawn from a general (normal) population of children.

(2) Schools (8/9 YO children) and Early Year Settings (3/4 YO children) were selected to ensure that the study samples reflected the full range of socio-economic backgrounds of age-matched children in the study area.

(3) To further assess how representative the sample was in terms of behavior, the teachers in the participating Early Year Settings and schools completed a behavioral profile questionnaire for all 3/4 YO and 8/9 YO children. The study sample was found to be representative of the general population of age and gender-matched children in terms of the teacher behavioral ratings.

(4) The selection procedure was structured such that for each age group there were no significant differences in socioeconomic or other background characteristics between any of the three sample groups used for statistical analyses or between groups of children assigned to receive the challenge drinks in different orders over the 6 week study period.

(5) The experimental study was appropriately designed as a within-subject crossover with three challenge treatments, two additive mixes (Mix A and Mix B) and a placebo administered in fruit juice under double-blind conditions.

(6) Multiple behavioral measures were used to detect any treatment effects. The measures included standardized parent ratings, teacher ratings, classroom /playroom observations, and a continuous performance task (8/9 YO children only).

(7) Several correlative secondary research questions were also addressed in this study, including: 1) whether challenge related behavioral effects are seen across the different sources of behavior measures: parent ratings, teacher ratings, direct classroom/playroom observation of behavior, and continuous performance testing (8/9YO only); 2) whether behavioral effects of AFCAs (Mix A and Mix B) may be modulated by genetic differences between children in a normal population; and 3) whether it is possible to demonstrate short term changes in behavior immediately after single dose acute challenge with a mix of AFCAs and to identify metabolic factors that may mediate such responses.

VIII. REVIEWER NOTED STUDY WEAKNESSES

There are a number of weaknesses that impact confidence in the study findings to various degrees. Specific weaknesses or shortcomings of this study are presented below in relation to (1) the procedures (Materials and Methods) used in the conduct of this study and (2) the analyses of the

study data and interpretation of the study findings.

A. <u>Procedural Weaknesses</u>

(1) There is questionable accuracy of the teacher and parent pre-study behavioral ratings. The teachers and parents completed behavioral questionnaires prior to the start of the study to indicate frequency of certain behaviors for individual children over the previous 6 months. The accuracy of ratings based on a recall of behaviors over a period of 6 months is highly questionable, particularly for teachers rating each of multiple children in a classroom or playroom setting.

(2) The two active challenges (A and B) used in this study were both mixtures of four color additives plus sodium benzoate. The use of chemical mixtures as challenge materials precludes identifying which specific compound(s) within the mixtures might be responsible for any treatment-related effects. In order to identify the specific active compound(s), subsequent studies would be needed in which each chemical is individually is used as a challenge substance, preferably at several dose levels.

(3) No explanations were given for the use of different color additives between Mix A and Mix B making the composition of the two mixes different; for the disproportionate increases in the levels of the color additives in the mixes (A and B) used for the 8/9YO children relative to the mixes (A and B) used for the 3/4YO children; and for the absence of adjusting the levels of sodium benzoate between age groups of children effectively lowering the dose level of sodium benzoate for the older children. Each of these procedural elements complicates the comparative assessment of treatment effects between Mix A and Mix B within or between age groups of children and limits the usefulness of such comparisons.

(4) The composition of the placebo was not clearly described. The placebo is a critical component of this, or any, challenge study and it would be important to have information about the types of fruit juice used (commercial or fresh). If commercial, what brands were used and did they contain any types of additives? If fresh, how was the juice made and when was it prepared relative to use in the study? Such information about the placebo source and composition are very important in helping to assess the outcome from this study.

(5) The parents were not asked whether they could tell the difference between the active and placebo drinks, as a check that the blinding was intact in the home environment.

(6) The detection of treatment related behavioral effects in the primary study was based principally on parental ratings, but the specific behavioral elements in the parent questionnaire that were most affected were not identified.

(7) The investigators did not provide comparative standard teacher/parent behavioral ratings or COC ratings for ADHD and non-ADHD children, with which to help gauge the significance of the ratings for the children in this study or to help identify responders versus non-responders.

(8) Each child's classroom observation score for an entire week was based on only three 8-minute observation periods (total of 24 minutes of observation per week). It is

questionable whether observing a child for only 8 minutes three times a week is an adequate sampling frequency or observation duration to provide a reliable measure of representative classroom behavior over an entire week.

(9) In the "acute challenge phase" of this study the disaggregated behavior measures comprising the Hyperactivity Index (HI) were not analyzed separately. The component behavior measures, CPT and independent behavior ratings, were aggregated to form the Hyperactivity Index (HI). However, these component behavior measures were not analyzed separately (disaggregated) to help determine whether changes in CPT performance or the behavior rating contributed more, less or equivalently to any treatment related changes in the overall aggregate HI scores.

(10) The Mixed Models analyses of the "acute challenge" data did not appear to control for potential confounding factors. In the absence of controlling for confounding variables there is low confidence in the reliability of the statistical analyses of the "acute challenge" study findings.

(11) No methodological details were provided for the genotyping analyses. The investigators did not provide any description of the materials used or any procedural details about the manner in which DNA samples were taken, at what period during the main study the samples were taken, how the samples were maintained, or how the samples were analyzed for SNPs. Without such methodological information it is difficult to assess the adequacy of the procedures used in this analysis or the reliability of the data collected.

(12) With reference to the secondary question of whether genotype may have a modulating influence on the behavioral effects of AFCA treatment, one important aspect that was not considered is the comparative genotype in the sub-set of AFCA "responders" versus "non-responders". Determining the relative distribution of the presence or absence of each of the polymorphic alleles for "responders" and "non-responders" would have complemented the determination of whether the presence or absence of certain gene polymorphisms correlate with the adverse effects of AFCA on behavior.

B. <u>Weaknesses Regarding Data Analyses and Interpretation of Study Findings</u>

(1) In the statistical analyses for treatment related changes in overall GHAs or in the component (disaggregated) behavior measures for both age groups of children, the finding of statistically significant effects varied notably depending upon whether the *whole sample*, the *≥85% consumption* sub-sample, or the *complete case* sub-sample of children was used for the analysis. For example, in the analysis of overall GHAs when either the *whole sample* or the *≥85% consumption* sub-sample of 8/9YO children was used for analysis, only Mix B had a significant effect. However, when the *complete case* sub-sample of 8/9YO children was used, both Mix A and Mix B had significant effects on GHA scores. In the analyses of the component (disaggregated) behavior measures, when the *whole sample* of children was used, there were no significant effects for the 3/4YO children and for the 8/9YO children only the parent ratings were significantly affected with Mix B challenge. None of the teacher ratings or classroom observation ratings (both age groups) or CPT scores (8/9YO) were significantly

affected. However, when the analyses used the *complete case* sub-sample of children, parental ratings were significantly affected only for the 3/4YO children with Mix A, and CPT scores for the 8/9 YO children were significantly affected only with Mix B. The teacher and classroom observation ratings were still not significantly affected. This variability in statistically significant outcomes may indicate that non-compliance in consuming the challenge drinks or the method used to handle missing data in the analyses had affected the pattern of results. But in describing the results the investigators did not discuss these problematic aspects of the data analyses, in particular what relevance they have to the interpretation of the study findings. This lessens confidence in the reliability of the statistically significant study findings.

(2) Regarding the decreasing CPT performance of the 8/9 YO children over the study period, little detailed information was presented, but the study investigators concluded that this decline was due to the children becoming bored with the repeated conduct of this inherently tedious task. Sufficient additional information should have been provided to support this conclusion.

(3) None of the confounding variables, which were controlled in the data analyses of the primary study, were incorporated into the Mixed Model methods used for analysis of the "acute challenge" data. Since potential confounds were not controlled in the data analysis, this calls into question the reliability of the findings from the "acute challenge" component study.

(4) The inconsistent finding of significant main effects of challenge (Mix A and Mix B) on behavioral responses (GHA) in the analyses of the genotype data raises questions about replicability of the Mixed Models analyses and uncertainties about the findings of significant treatment effects in the primary study. In each age group of children (3/4YO and 8/9YO), to test for the influence of four genotypes on the behavioral effects of AFCAs, a Mixed Models analysis was conducted on the data of four sets of subjects, taken only from the *≥85% consumption* sub-sample, with each set comprised of virtually the same subjects. Since the analysis of each of the four genotypes included a test for main effect of challenge (Mix A and Mix B) on GHA, the analysis for main challenge effects was repeated four times using basically the same group of subjects. If a statistically significant main challenge effect of treatment was present, it would have been expected to be found for each of the four sets of subjects. While a significant main challenge effect was present, this effect was not consistently detected across the four sets of subjects. In the 3/4YO children main challenge effects for both Mix A and B were found in only two of the four analyses; in the 8/9YO children significant main effects were found for Mix A in two of the four analyses and for Mix B in three of the four analyses. This inconsistency in finding significant main challenge effects raises questions regarding the replicability of the Mixed Models statistical analysis. Since the Mixed Models analysis was used to analyze the primary study data, there are uncertainties about the replicability and the relevance of the statistically significant challenge effects in the primary study. A specific evaluation by a statistician of the adequacy and replicability of the statistical procedures used for the primary study and the genotype analyses would be appropriate.

(5) The rationale offered by the investigators to explain why treatment effects in the primary study appeared to be based primarily on the parental ratings with little

contribution from the other behavior measures was based on an inaccurate interpretation of the acute study findings. The principle investigators contend that the findings from the acute challenge study suggest that the 8/9 YO children who respond to the ingestion of Mix B challenge do so within a short period of time, an hour. Since the children in the primary study consumed their challenge drinks at home, the findings from the acute challenge study would suggest that it was more likely for the parents in the home setting, rather than teachers or observers, to have observed to the behavioral changes and may be an explanation as to why the effects were detected principally by the parental ratings rather than the school or early years settings measures. The trends in the acute challenge study (although not statistically significant) may suggest that the onset of response following ingestion of Mix B challenge occurred within a short period of time, an hour. However, the acute challenge study is not designed to determine the duration of any treatment effects. Consequently, though some of the children may have consumed their challenge drinks at home after returning from school (or from the Early Year Setting) and experienced the onset of treatment effects shortly thereafter, the acute study provides no information about how long those effects may have lasted. It is possible that treatment effects, if any occurred, could have persisted at least into the day after challenge, when the children returned to the school or the Early Year Setting environments.

(6) The use of the three sample groups of children for analysis of the component (disaggregated) behavior measures resulted in variable statistical outcomes but relevance of this to data interpretation was not discussed in the technical report. As was apparent in the analyses of the change in GHAs for the component behavior measures, there is notable variability in the occurrence of statistically significant effects across subgroups. Depending upon whether the *whole sample*, the *≥85% consumption* sub-sample or the *complete case* sub-sample of children was used for the analysis, the outcome of significant effects on certain component behaviors is notably different. For example, based on analysis of the component behaviors using the *whole sample* of children, there were no significant effects for the 3/4YO children, and for the 8/9YO children only the parent ratings were significantly affected with Mix B challenge. None of the teacher ratings or classroom observation ratings for either age groups or CPT scores (8/9YO) were significantly affected. However, when the analyses used the *complete case* sub-sample of children, parent ratings were significantly affected only for the 3/4YO children with Mix A, and CPT scores were significantly affected but only for the 8/9YO children with Mix B. The teacher and classroom observation ratings were still not significantly affected. Although this obvious variability in significant treatment effects across subgroups of subjects may indicate that non-compliance in consumption of challenge drinks or the method of handling missing data affected the pattern of results, the investigators do not discuss the impact of this problematic aspect of the data analyses on interpretation of the study findings.

(7) There was no basis presented for considering the behavioral changes in this study as necessarily "adverse" or "deleterious". In the previous study by Bateman et al (2004) treatment effects for Mix A were found based on the daily parental ratings, but were not confirmed in the weekly clinical assessments by research psychologists using validated tests. In the present study the *whole sample* analyses showed significant effects of Mix A on GHA levels for the 3/4YO children and significant effects of Mix B on GHA for the 8/9YO children. Analysis of the disaggregated component behavior measures for

the *whole sample* of subjects (see Results section) indicated that parental ratings for both age groups of children and CPT scores for the 8/9YO children (not measured in the younger children) were the major contributors to the significant AFCA effects on the overall GHA levels. Notably, there appeared to be little suggestion of treatment effects based on either the teacher ratings or the classroom observation ratings, indicating that Mix A and Mix B did not elicit behavioral changes that adversely affected the early years setting (3/4YO children), school (8/9YO children) or observation classroom (both age groups) environments. There was no information provided in either study to suggest that the changes in behavior based on parental reports were adverse, detrimental or maladaptive. The children's behavior under challenge conditions appeared to be within the range of behavioral levels exhibited by the general population of age-matched children.

(8) The suggested utility of genotype results from this study of behavior in a general population of children to studies dealing with the specific condition of ADHD is an inappropriate extrapolation. The present study attempts to provide some information related to whether any effects of AFCAs on behavior ('*hyperactivity profile behaviors'* of over activity, inattention and impulsivity) are moderated by genetic differences between children in a normal population. This is not the same as a population of ADHD children. ADHD is a specific neurological disorder characterized by dysfunctional behavior and is not the extreme end of a biological continuum with normal behavior.

IX. Applicability to Assess Risk or to Support Regulatory Action

The primary findings in the present study showed that 1 week of daily challenges with certain mixtures of artificial food colors and sodium benzoate may produce low level behavior effects in 3/4 YO and 8/9 YO children from the general population, limited to detection based primarily on parental ratings and possibly continuous performance test scores. There was no indication that any of these behavioral findings constituted clear adverse effects. However, due to the absence of confirmation of the parental rating effects by teacher ratings and classroom observation ratings, along with concerns about data analyses and various procedural weaknesses, there is questionable confidence in the reliability and biological relevance of the primary findings from this study. One particular procedural weakness relevant to regulatory application was the use of chemical mixtures as challenge materials which precludes identifying which specific compound(s) within the mixtures might be responsible for any treatment related effects. Consequently, there would be little, if any utility of these findings to assess risk or to support any specific regulatory decision.

Preliminary genotype findings, which were developed in this study to address a secondary question, tended to support the possibility of some level of genetic influence on the sensitivity to the behavioral effects of certain AFCAs in children from the general population. However, due to data analysis and procedural concerns there are uncertainties regarding the adequacy of the statistical and analytical procedures used and the reliability of the data collected. These uncertainties, together with the preliminary nature of these genotype findings, would indicate little, if any, regulatory utility for these findings.

References

Angello L, Volpe R, DiPerna J, Gureasko-Moore S, Gureasko-Moore D, Nebrig M and Ota K. An assessment of attention-deficit/hyperactivity disorder: An evaluation of six published rating scales. School Psychology Review 2003; vol. 32.

Bateman B, Warner JO, Hutchinson E, et al. The effects of a double-blind, placebo controlled, artificial food colourings and benzoate preservative challenge on hyperactivity in a general population sample of preschool children. Arch Dis Child 2004; 89: 506-11.

NIH Consensus Conference (Editorial). NIH consensus development conference: defined diets and childhood hyperactivity. Clinical Paediatrics 1982; 21: 627-30.

Feingold BF. Hyperkinesis and learning disabilities linked to artificial food flavors and colours. Am J Nurs 1975; 75: 797-803.

Schab DW, Trinh NT. Do artificial food colours promote hyperactivity in children with hyperactive syndromes? A meta-analysis of double-blind placebo-controlled trials. J Dev Behav Pediatr 2004; 25: 423-34.

Swanson JM, Sergeant J, Taylor E, Sonuga-Barke E, Jensen PS, Cantwell D. Attention Deficit Hyperactivity Disorder and Hyperkinetic Disorder. The Lancet 1998; 351: 429-33.

TECHNICAL DATA EVALUATION REPORT

Proposed Association Between
Artificial Food Colors and Attention Deficit Hyperactivity Disorders
(ADHD) and Problem Behaviors in Children

Prepared for
Center for Food Safety and Applied Nutrition'
Office of Food Additive Safety
U.S. Food and Drug Administration
College Park, MD 20740-3835

Prepared by
Toxicology and Hazard Assessment Group
Environmental Sciences Division
Oak Ridge National Laboratory
Oak Ridge, TN 37831
Task 2008-30

Primary Reviewer:
Thomas J. Sobotka, Ph.D.

Signature:
Date: NOV 2 9 2010

Secondary Reviewers:
Robert H. Ross, M.S., Group Leader

Signature:
Date: NOV 2 9 2010

Oak Ridge National Laboratory managed and operated by UT-Battelle, LLC., for the U.S. Department of Energy under Contract No. DE-AC05-00OR22725

THE FINDINGS AND CONCLUSIONS IN THIS REPORT HAVE NOT BEEN FORMALLY DISSEMINATED BY FDA AND SHOULD NOT BE CONSTRUED TO REPRESENT ANY AGENCY DETERMINATION OR POLICY.

Addendum Review to the ORNL IAG Project Assignment (#2008-31) on "*The Evaluation of proposed Association Between Artificial Food Colors and Attention Deficit Hyperactivity Disorders (ADHD) and Problem Behaviors in Children*"

Thomas J. Sobotka, Ph.D.
August 25, 2010

Table of Contents

I INTRODUCTION 2
II REVIEW and EVALUATION OF *SCHAB and TRINH (2004)* 2
A. Study Rationale and Design 2
B. Study Results and Conclusions 3
C. Assessment and Summary Evaluation of Study 5

I INTRODUCTION

The present review and evaluation report serves as an addendum to the ORNL IAG project assignment (#2008-31) on "*The Evaluation of proposed Association Between Artificial Food Colors and Attention Deficit Hyperactivity Disorders (ADHD) and Problem Behaviors in Children*". The subject of this review/evaluation is a study conducted by *Schab and Trinh (2004*) which presents the results of a meta-analysis of findings from previously conducted clinical trials addressing the proposed relationship between consumption of artificial food colors (AFCs) and behavioral changes in children diagnosed with hyperactivity (ADHD). This meta-analysis has been cited by the Center for Science in the Public Interest (CSPI; Petition submitted to the FDA on June 3, 2008; File FDA.2008.!0349.0001) as a pivotal study in CSPI's assertion that certified color additives cause adverse behavioral effects in some children; therefore, it was deemed important that a thorough evaluation of this study should be included in FDA's review of the literature.

II REVIEW and EVALUATION OF *SCHAB and TRINH (2004)*

Schab DW Trinh NT. Do artificial food colors promote hyperactivity in children with hyperactive syndromes? A meta-analysis of double-blind placebo-controlled trials. J Dev Behav Pediatr. 2004 (December), 25(6):423-34.

A. Study Rationale and Design

The study conducted by Schab and Trinh (2004) attempted to provide a focused evaluation of whether artificial food colors (AFC) promote the symptomatology of hyperactivity in children diagnosed with hyperactive syndromes, as measured on behavioral rating scales. These authors conducted a primary meta-analysis of findings, based only on behavior ratings, from 15 selected previously conducted double-blind placebo-controlled clinical trials (total of 219 subjects) addressing a proposed relationship between consumption of AFCs and behavioral changes in children (7.9 years average age; male-female ratio at 5.5 to 1) with a diagnosis of hyperactivity. It should be emphasized that, other than the behavioral ratings (parent, teacher, clinician), this meta-analysis did not consider any of the more objective outcome measures of behavior such as clinical/psychological evaluations, activity monitoring, and behavioral testing which were available in many of the trials. The same study data as used in the primary meta-analysis were also analyzed to address three sub-hypotheses that may help explain the heterogeneity of prior trials' results, facilitate identification of responders to AFCs, or aid in design of future trials: (1) *Sub-hypothesis 1*: parents and teachers differ in their reports of responsiveness to AFCs; (2) *Sub-hypothesis 2:* open trial and parental report can be used as screening methods to identify potentially responsive children; and (3) *Sub-hypothesis 3:* children formally diagnosed with hyperactivity have greater responsiveness to AFCs than do heterogeneous or non-hyperactive groups of children. Additionally, a secondary analysis was conducted to address the three sub-hypotheses using data from an additional 8 clinical trials (total of 132 subjects) which met the primary inclusion criteria except that the subject population consisted of either exclusively non-hyperactive or a heterogeneous mixture of hyperactive and non-hyperactive children (7.3 years average age; male-female ratio at 3.5 to 1). To standardize the behavior rating data across clinical trials, the authors used a random effects model with the standardized mean difference

(SMD) as the measure of effect size (ES). The SMD describes the difference in outcome between the active and control arms of a trial in terms of the number of pooled standard deviations by which the two groups differ. The summary ES depicts the summary response to AFCs. Adverse effects on hyperactivity are reflected in positive ES values.

B. Study Results and Conclusions

Primary Analysis w Hyperactive Children:

Based on the behavioral ratings of parents, teachers and clinicians combined from 15 clinical trials meeting the primary inclusion criteria with children diagnosed with hyperactivity syndrome, meta-analytic modeling identified a statistically significant overall summary response to artificial food colors (AFCs) with an effects size (ES) of 0.283 (95% CI, 0.079 to 0.488), reflecting a change of slightly more than 1/4 of a SD. The authors state that the magnitude of this ES is approximately 1/3 to 1/2 the magnitude of the ES calculated in a meta-analysis of trials evaluating methylphenidate as a treatment of ADHD (Schachter et al, 2001). The authors conclude that these results suggest an association between ingestion of AFCs and problem behaviors in hyperactive children. The need for qualifications to this conclusion is explained below in the section "Assessment and Evaluation of the Study".

Testing the *sub-hypothesis 1* showed that, while the ES values for teachers (ES = 0.0810; 95% CI, -0.073 to 0.235) and clinicians (ES = 0.107; 95% CI, -0.128 to 0.343) were not statistically significant, the ES values for parents (ES = 0.441; 95% CI, 0.161 to 0.721) were statistically significant. However, since no statistical difference was found between these three groups' scores and no relevant information is provided in the secondary analysis, the authors concluded that the *first sub-hypothesis*, that parents, teachers and clinicians differ in their reports of hyperactivity behaviors, could not be confirmed. Although this latter conclusion may be statistically correct, it is practically incongruous with the fact that, as the authors state, "only the parents' summary ES was statistically significant". Effectively, as the authors also state, "in our results, parents tended to detect greater change from the intervention [AFCs] than did teachers [or clinicians]". As discussed below in the section "Assessment and Evaluation of the Study", this issue of parents detecting behavioral changes not noticed by other raters is relevant to interpretation of the results from this meta-analysis.

Testing the effectiveness of screening hyperactive children as potentially responsive to AFCs (*sub-hypothesis 2*) revealed that trials that did not prescreen for potential responders demonstrated a nonsignificant effect of AFCs on behavior ratings in controlled testing (ES = 0.09; 95% CI, -0.108 to 0.288) but trials that selected subjects who were prescreened as presumptively responsive to AFCs did demonstrate a statistically significant effect of AFCs (ES = 0.535; 95% CI, 0.149 to 0.920). However, there was no statistically significant difference between these two groups. The significance of these findings will be discussed in conjunction with the results from the secondary analysis below.

As a partial approach toward addressing the *third sub-hypothesis* that hyperactive children have greater responsiveness to AFCs than non-hyperactive children or children

heterogeneous for hyperactivity, the trials in the primary analysis which used only hyperactive children were ordered according to diagnostic grades, i.e. severity of hyperactivity. No differential trends in ES values were found across these groupings of trials. The significance of these findings will be discussed in conjunction with the results from the secondary analysis below.

Secondary Analysis w Non-Hyperactive or Heterogeneous Groups of Children: Based on 8 clinical trials (secondary trials) meeting the primary inclusion criteria with non-hyperactive children or groups of children heterogeneous for hyperactivity, meta-analysis revealed a non-significant ES of 0.117 (95% CI, -0.113 to 0.347) indicating no summary effect of AFCs on behavior ratings for this group of trials.

In the absence of a significant ES, the secondary trials were not tested for *sub-hypothesis 1* (i.e., difference between parents' and teachers' ratings).

However, when the secondary trials were analyzed separately for those that pre-screened or did not pre-screen for potential responsiveness to colors (*sub-hypothesis 2*), there was a significant difference between these two groups. Trials that did not pre-screen showed a non-significant response to AFCs in controlled testing (ES = -0.112; 95% CI, -0.393 to 0.169); however, a statistically significant ES of 0.316 (95% CI, 0.157 to 0.475) characterized the trials that did pre-screen. The difference in ESs between these secondary trials that prescreened and did not prescreen children was statistically significant. Post-hoc analysis combining the primary and secondary trials also showed that, among all trials, those that prescreened were significantly different from those that did not screen. Based on both the primary and secondary analyses, the authors concluded that the *second hypothesis*, that children potentially responsive to AFCs can be identified through screening methods, was only partially substantiated. Overall, trials that pre-screened for responsiveness demonstrated a significant treatment effect in controlled testing whether enrolling hyperactive (primary trials) or heterogeneous, predominantly non-hyperactive subjects. However, there was a statistically significant difference between trials that pre-screened and those that did not only among the secondary group of trials enrolling heterogeneous, predominantly non-hyperactive subjects. Interestingly, the authors suggested that "screening may be more effective in heterogeneous populations because of greater variability in responsiveness to AFCs among such populations than among the strictly hyperactive populations". This reviewer suggests the possibility that this unique responsiveness to AFCs may be associated with a genetic polymorphism that may be more variable in the heterogeneous populations than in the strictly hyperactive populations (refer to Color/ADHD Overview Report by Sobotka, 2010).

In testing the *third sub-hypothesis*, that hyperactive children have greater responsiveness to AFCs than non-hyperactive children or children heterogeneous for hyperactivity, the summary ES of the trials in the secondary analysis, whose participants were predominantly non-hyperactive (but with other problem behaviors), did not differ significantly from the summary ES of the trials in the primary analysis (hyperactive children). The authors concluded that this negates the *third sub-hypothesis* and shows that presence of a baseline diagnosis of hyperactivity and rigor of that diagnosis does not correspond to reactivity to AFCs, suggesting that AFC-sensitivity among patients may not be limited to those with clear-cut criteria-specific hyperactive syndrome.

C. Assessment and Summary Evaluation of Study

One prefacing comment relevant to assessing this study and its conclusions is that one important element in the conduct of a meta-analysis involves consideration of the quality of the various trials included in the analysis and the confidence in the data and findings from those trials. According to the authors of this study, the various trials used in this meta-analysis were "graded according to the degree that they adequately addressed types of bias and avoided important threats to validity: differential selection, performance, attrition, detection, and reporting". In the present study (page 426, section "*Primary Analysis*") the various trials were assigned "validity" scores of "A" (the highest score) to "C" (the lowest). Although the scoring criteria were not specified and the individual scores for all trials were not presented, among the 15 trials used for the Primary Analysis, 2 trials received the highest "A" validity score, 2 trials received the lowest "C" score, and apparently most of the trials (11) were scored as having a moderate "B" level of validity. It is unclear, however, how or whether the weighting values of these "validity" scores and other experimental limitations were actually factored into the basic primary meta-analysis in this study. The only described use of the "validity" scores, as indicated on page 428 of the document (section "*Sensitivity Analysis of Primary Trials"*), was in the conduct of an additional separate meta-analyses which excluded 2 trials with "C" validity scores. This separate analysis reduced the ES (Effect Size) from the summary primary meta-analysis of 0.283 (statistically significant) to a recalculated ES of 0.216 (statistically significant). An additional separate meta-analysis was conducted to exclude 2 trials with the smallest numbers of subjects from the analysis, reducing the ES score to 0.255 (statistically significant). Analysis with both types of trials excluded (lowest quality and smallest) yielded a summary ES score of 0.210 (still statistically significant). Since there is no other mention of the "validity" scores being used in this analysis, it appears, therefore, that the primary meta-analysis (ES = 0.283) in this study was conducted with findings from the various trials that were used with little adjustment of the analysis to factor in measure(s) of quality of the clinical trials and confidence in the reported findings from those trials. However, in the Discussion section of this paper the authors do attempt to explain the significance of many of the experimental design issues and other shortcomings of the trials used for this meta-analysis and their impact on interpretation of this study's results. The authors also include recommendations whereby future research may avoid the various problematic experimental methods and design issues which many prior trials have experienced. This reviewer suggests that input from an experienced statistician would be beneficial in providing a more cogent assessment of the appropriateness of the statistical procedures used in the conduct of this meta-analysis, including the significance of the "validity" scores in accounting for the quality of the trials and confidence in their findings.

The authors open their discussion of their study's findings with the statement that this "meta-analysis supports the hypothesis that AFCs promote hyperactivity in hyperactive children, as measured on behavioral rating scales". This reviewer considers this a rather broad overstatement of the significance of their findings. Later in the discussion the authors do state appropriately that "several limitations, especially of the included trials, but also of our meta-analytic methods, restrict the interpretation of our results" and conclude more correctly that their results only "strongly suggest an association of AFCs

and hyperactivity". Among the limitations referred to by the authors were the use of unorthodox scales to measure behavioral responses which may limit the validity of conclusions about the effects of AFCs on hyperactivity (but suggest that AFCs provoke general behavioral disturbance more than hyperactivity), inappropriate statistical methods in many of the trials, clinical heterogeneity of the study population together with the trials' inconsistent reporting of subtypes, and the possibility of that this study's use of only behavioral rating outcomes may make this study irrelevant to ADHD's inattentive subtype, whose symptoms may not be predominantly behavioral phenomena. The latter limitation, that is the use of only behavioral rating outcomes for this meta-analysis, has a broader relevance to the interpretation of the results of this study particularly in view of the results of testing the *second sub-hypothesis* in this study that compared the sources of the behavioral ratings, i.e. parents, teachers, and clinicians. The outcome of testing the *second sub-hypothesis* showed that only the summary ES of parents' ratings was statistically significant but neither the teacher based nor clinician based ESs were significant (although there was no statistical difference between the three groups' scores). This appears to indicate that the parental behavioral ratings may be the major contributor to the overall significant ES associated with AFCs and that the teacher and clinician ratings provide little if any support for a behavioral response to AFCs. Several trials have reported similar findings with parent ratings showing effects associated with AFCs but teacher or clinician ratings providing contradictory or otherwise no supporting confirmation (Harley et al, 1979; Mattes and Gittelman-Klein, 1981; Williams et al, 1978). In discussing the *second sub-hypothesis* finding that parents may have detected behavioral change unnoticed by teachers and clinicians, the authors considered several possible reasons for this difference, including the following:

1. *that parents were ineffectively blinded* – The authors cited 5 trials (Goyette et al, 1978a and b; Harley et al, 1978; Mattes and Gittelman-Klein, 1978; Rose, 1977) as having conducted rigorous testing and confirmed the effectiveness of the blind for parents, subject and investigators (2 of the cited trials [Goyette et al, 1978 a and b] in fact provided no description of the blinding procedures used in their trials). This reviewer considers this less than convincing evidence to conclude that the parents were effectively blinded in the majority of trials used for the meta-analysis and to dismiss the possibility that ineffective blinding may have unintentionally resulted in the parents reporting behavioral changes that were not evident to either the teachers or clinicians. The authors even include "explicitly testing the blinding of subjects" as one of the recommendations whereby future research may avoid the various problematic experimental methods and design issues which many prior trials have experienced.
2. *that parents' based their ratings on problem behaviors of concern to them (irritability and sleep disturbance) which may differ from those of concern to teachers and clinicians (restlessness and inattention) and that AFCs may promote a pattern of symptoms that is incongruent with modern criteria of ADHD but that is nevertheless bothersome to parents* – this reflects suggestions made by several investigators the effects of AFCs may not be associated with ADHD types of behavior but more with general problem behaviors (Carter et al, 1993; Mattes and Gittelman-Klein, 1978; Rowe, 1988; Rowe and Rowe, 1994; Sarantinos et al, 1990).
3. *parents may be particularly attuned to the idiosyncrasies of their own children* – several investigators attempted to address this possibility by using specially developed questionnaires specifically compiled by parents (Weiss et al, 1980) or

investigators (Rowe, 1988) to target the unique behaviors of the subject children. Neither study identified a higher percent of responders than trials using standard rating questionnaires indicating that this is not a likely reason for parents to detect behavioral changes not evident to other raters.

In the absence of a suitable explanation for the detection differences between parent ratings and teacher and clinician ratings, this reviewer considers that parental reports of behavioral changes unconfirmed or not support by other sources or measures of behavioral change should generally be viewed as inconclusive and only suggestive of a treatment effect.

Overall, the primary results of this meta-analysis of behavioral rating findings from 15 controlled clinical trials with diagnosed hyperactive children did show a small but significant summary effect of AFCs on parental ratings, but no significant treatment effect based on either teacher or clinician ratings. In view of the absence of confirmation of treatment effects with either the teacher or clinician behavioral ratings and the questionable use of the “validity” scores for the trials used in this analysis, the significance of the effects identified in this meta-analysis is viewed as inconclusive but suggestive of an association between AFCs and hyperactivity behaviors. This interpretation may be considered consistent with the fact that the findings from the various clinical trials used in this analysis were generally considered only suggestive of a limited association between AFCs and adverse behaviors in children with hyperactivity or other problem behaviors. Importantly, the findings from this meta-analysis do not negate either the general view that only a small subgroup of children appears to be affected by AFCs or the suggestion that these responsive children may be among a subgroup of children who are uniquely sensitive (intolerant) to a variety of food related chemicals not limited to AFCs which may adversely affect behavior in this select subgroup of children. Additionally, the findings from this study do highlight questions other investigators have raised about the specificity or nature of effects associated with AFCs, specifically that the symptomatology of ADHD may differ from the pattern of symptoms induced by AFCs, the latter being associated more with provoking general behavioral disturbances such as irritability and insomnia than with eliciting ADHD symptoms such as restlessness and inattention (Carter et al, 1993; Mattes and Gittelman-Klein, 1978; Rowe, 1988; Rowe and Rowe, 1994; Sarantinos et al, 1990). The present authors suggest a broader focus such that "researchers should not limit their search for responders to those previously diagnosed with hyperactivity" and thereby may progress "out of the realm of hyperactivity and into a wider province of behavioral pathology".

THE FINDINGS AND CONCLUSIONS IN THIS REPORT HAVE NOT BEEN FORMALLY DISSEMINATED BY FDA AND SHOULD NOT BE CONSTRUED TO REPRESENT ANY AGENCY DETERMINATION OR POLICY.

Overview and Evaluation of *Proposed Association Between Artificial Food Colors and Attention Deficit Hyperactivity Disorders (ADHD) and Problem Behaviors in Children*

Thomas J. Sobotka, Ph.D.

June 13, 2010
(revised August 23, 2010)

Table of Contents

LIST of ATTACHMENTS 2
I Introduction 3
II Background 3
III Criteria for Review and Assessment of Clinical Trials 4
IV Summary Reviews/Assessments of Clinical Trials 6
Group I Trials (n=26) 7
Group II Trials (n=7) 27
V Overview of Animal/Laboratory Studies 35
VI Possible Biological Processes/Mechanisms 37
VII Overall Evaluation & Interpretation of Available Information 39
VIII Identification of Information Gaps/Suggested Additional Study 42
IX Conclusion 44
X Comment 44

LIST of ATTACHMENTS

- **Exhibit 1:** REVIEWS and CRITIQUES
- **Exhibit 2:** BIBLIOGRAPHY - ALL REFERENCES RELATED TO OVERVIEW ASSESSMENT
- **Exhibit 3:** Table 1 - SUMMARY of REVIEWED CLINICAL TRIALS

I Introduction

The present report addresses the project assignment (#2008-31) to conduct a critical review and analysis of peer-reviewed clinical trials, identified through a comprehensive literature search (ORNL), related to the consumption of food additives and problem behaviors in children, such as the attention deficit hyperactivity disorders (ADHD). This review focused on those trials that related to artificial food colors, since food color was the type of additive most commonly investigated in relation to possible adverse effects on children's behaviors, such as ADHD, and was also the basic focus of a citizen's petition to the FDA (dated June 3, 2008). However, many of the trials reviewed in this report also involved artificial flavors, natural salicylates, preservatives, other additives, and general food items in conjunction with artificial colors.

As requested in the project assignment, a thorough search of the literature from 1982 to the present was conducted and in-depth reviews of the identified relevant clinical publications (17 clinical trials) were carried out. In addition, 16 other relevant publications of clinical trials conducted prior to 1982 (between 1976 and 1981) were identified, most of which were also referenced in the citizen's petition to the FDA. For purposes of completeness and continuity, these pre-1982 clinical trials were also included in this review.

As noted in the attached *Bibliography*, this report was based on an in-depth review and analysis of a total of 33 clinical trials with supplemental information obtained from 17 animal/laboratory studies and 51 background documents, including various reports, commentaries, reviews, and meta-analyses. A summarized outline of the 33 clinical trials in chronological order is presented in *Table 1* (the 16 pre-1982 clinical trials are identified by grey highlighting of the Author/Year in column one of the Table) and the detailed review of each is presented in the attached *Reviews and Critiques* section (and file) of this report.

II Background

Among the variety of problematic behaviors in children that are particularly disruptive in the home and school environment is a spectrum of behavioral disorders characterized by a pattern of behavior deficits including inattention, impulsivity, disinhibited behavior, and overactivity. Of course, similar types of behaviors are seen in most children in the general population (Conners, 1970a: *Conners, 1970b; Pollock, 1991; Stevenson, 2006; Taylor, 1984*), but when these behaviors are determined by professional judgment to occur in a developmentally and situationally inappropriate manner, persist over a prolonged period of time and at a high level of severity, and possibly be associated with learning disabilities, they may be considered part of a pathological state of hyperactivity (Pollock, 1991), commonly referred to as attention deficit hyperactivity disorder of childhood (ADHD) (*Diagnostic and Statistical Manual III* of the American Psychiatric Association). Other terms used include hyperactivity syndrome of childhood, hyperkinetic syndrome, hyperkinesis, minimal brain dysfunction, attention deficit disorder, and conduct disorder, although some distinctions may exist between these terms (NIH Consensus, 1983; Pollock, 1991; Ribon and Joshi, 1982; Schab and Trinh, 2004; Schauss, 1984; Silbergeld and Anderson, 1982; Taylor, 1984). ADHD is generally estimated to occur in approximately 3 – 10% of the population of school children (Banergee et al, 2007; National Institutes of Health, 1998) with a male preponderance of about 4:1 (Pollock, 1991).

Many possible factors have been suggested to be involved, singly or synergistically, in the etiology of childhood ADHD syndromes, including environmental, genetic, allergic/immunologic, psychosocial, and dietary (Arnold, 1999; Banerjee et al, 2007; Cormier and Elder, 2007; Cruz and Bahna, 2006; NIH, 1982; Pollock, 1991; Stevenson, 2006;Taylor, 1984; Wender, 1986). Focused interest in the association between diet and hyperactivity and other problem behaviors in children was stimulated in 1973, when an hypothesis was put forward that certain food additives, specifically artificial food colors and flavors and natural salicylates, may cause or exaggerate a number of childhood behavior disorders, such as hyperkinesis (ADHD) and learning disabilities (Feingold, 1973 and 1975). Based on clinical observations, Feingold claimed that some children with problem behaviors, including hyperactivity and learning disabilities, showed notable improvement when given a defined diet excluding foods with artificial colors, flavors and natural salicylates and that dramatic deterioration of behavior occurred when the children were exposed to foods with even small amounts of those additives, particularly the artificial colors. Feingold's elimination diet rapidly became a popularized treatment for hyperactivity among the general public, even though the scientific community continued to question the validity of Feingold's uncontrolled anecdotal clinical observations (Wender, 1986). The need for sound scientific study to assess the validity of Feingold's claims that foods with artificial colors and other additives can cause or trigger ADHD types of problem behaviors in susceptible children was evident and resulted in a series of peer-reviewed clinical trials being conducted over the last 35 years, along with a vast array of related published reports, reviews, analyses, and commentaries. A total of 31 clinical trials, fairly representative of those conducted over this period of time and considered relevant to an assessment of the proposed association between artificial food colors and problem behaviors in children, were included in this project review.

III Criteria for Review and Assessment of Clinical Trials

In the review and assessment of each clinical trial, it was deemed important to identify all study findings, both positive and negative, and to consider any relevant experimental factors, particularly limitations, and inconsistencies or uncertainties in the data that may impact on the credibility, reliability and interpretability of the reported study findings. Further, in the interpretation of results efforts were made to distinguish between statistically significant and clinically relevant findings since some outcome measures may reveal statistical differences between treatment and control that may be within normally accepted standards. To this end, the general criteria used for assessing the trials and interpreting the findings included due consideration of the following (Krummel et al, 1996; NIH, 1982; Schab and Trinh, 2004; Wender, 1986):

- homogeneity of sample
- randomization to treatment
- crossover designs with subjects serving as own control
- counterbalanced treatment/challenge order
- double-blind/placebo-controlled challenges
- placebo and challenge indistinguishable
- *verification of effectiveness of blinding particularly for behavioral raters
- appropriate control outcome measurements
- age-appropriate outcome measures
- use validated measures (i.e. detect behavior differences/sensitive to treatment)

- *confirmatory sources of outcome data (parents, teachers, testing, etc.)

* Two of the above factors were considered by this reviewer to be particularly important in the evaluation and assessment of the study findings, specifically *confirmatory sources of outcome data* and *verification of effectiveness of blinding*. The importance of the factor *confirmatory sources of outcome data* was related to determining how much weight should be given to this factor in assessing reliability of study findings? Is a positive effect based on one source of data sufficient to conclude a reliable finding? The difficulty in making this determination was compounded by the fact that some trials used only one source of data, while other trials used multiple sources. Out of 33 clinical trials, 11 used only a single source of outcome data to assess treatment related effects on behavior and 8 of those showed positive effects; no other source of data was included in those studies to provide any confirmation of the positive findings. The remaining 22 clinical trials used multiple sources of outcome data but in 8 of those trials only one of the multiple sources of data showed positive treatment effects and no confirmation or support of a treatment effect was provided by the other sources of data in each of those 8 trials. Lipton and Mayo (1983) considered treatment related effects detected by one outcome measure (i.e., parent rating) to be suspect if there was no confirmation by other observers or objective tests. Similarly, Stevenson (2007) considered the study results reported by Batemen et al (2004) as somewhat equivocal, since they only related to the parent ratings. In making the determination of weighting and level of confidence in the present review, this reviewer adopted a conservative approach and in most cases considered any findings that were based solely on one source of data to be suggestive, but not conclusive evidence, of a treatment effect. Less confidence was given to the reliability of a single source finding when other sources of assessing behavioral change were used in the same study and provided data showing no confirmation or supportive evidence of any treatment related effects. It should be noted that many of the trials (14/33) cited in the present review reported treatment related effects based only on a parental rating outcome measure (either as the only source outcome measure used or the only outcome measure to detect a treatment effect with no confirmation by other behavioral measures). Consequently, since only a single source outcome measure detected an effect, a lowered weighting/level of confidence was generally assigned to these study findings. However, it should be clear that this lower weighting/less confidence was not selectively directed at the use of parental ratings but at the use of a single source outcome measure of behavioral change. Parental ratings under controlled clinical conditions are viewed as uniquely sensitive and as providing relevant information in the assessment of treatment related changes in behavior (Schab and Trinh, 2004), although it is considered preferable if they could be supported with observations by independent observers or the use of standardized tests (Stevenson, 2007). The main issue in assigning the weightings/confidence levels in this review was confirmation of a treatment related finding across outcome measures within the same study.

The importance of *verifying the effectiveness of the blind* stems from the widely acknowledged tenet that in any clinical therapeutic study the use of appropriate blinding procedures is essential to preclude placebo effects and observer bias from influencing the study results. In non-blinded studies the participants are typically aware of the subjects' treatment conditions and, consequently, the possible influence of placebo effects and observer bias lessens confidence in the reliability of those study findings. Double-blind studies, however, are considered more reliable, since various efforts are made to prevent (blind) the researchers, observers and subjects from knowing when the active or placebo treatments are administered (CSPI, 2008). However, even in double-blind studies, it is possible that the blinding procedures used in a particular study

may not be completely adequate or effective. Therefore, it is important that the effectiveness of the blinding procedures be assessed in each study. Without such assurance of the effectiveness of the blinding, the reliability of the study's findings may be considered questionable.

IV Summary Reviews/Assessments of Clinical Trials

An in-depth review and analysis for each of the 33 clinical trials included in this report is contained in the attachment, *Reviews and Critiques*. Those reviews are presented in that attachment in the chronological order in which they were published. This section of the report will present a brief summary and assessment of the findings from each of the clinical trials. *Table 1* is provided as a reference outline of each trial (in chronological order), color-coded† to identify those trials that used the Feingold diet, other diets, color challenges or food item challenges as the experimental treatment and to indicate for each trial whether treatment effects were found for each outcome measure used in that study.

The 33 clinical trials reviewed in this report are separated into two general groups based on a notable difference in one particular feature of the experimental design, that is, the nature of the experimental treatment. Group I consists of 26 clinical trials (color-coded pink† in Table 1) that are characterized by the use of an experimental treatment consisting of either the Feingold diet (eliminating foods with artificial colors, flavors and natural salicylates) or challenges with select artificial food colors (Conners et al, 1976; Harley et al, 1978 PI; Harley et al, 1978 PII; Goyette et al, 1978; Rose, 1978; Levy and Hobbes, 1978; Levy et al, 1978; Mattes and Gittelman-Klein, 1978; Williams et al, 1978; Conners, 1980; Conners et al, 1980; Swanson and Kinsbourne, 1980; Weiss et al, 1980; Mattes and Gittelman-Klein, 1981; Adams, 1981; Spring et al, 1981;Salamy et al, 1982; Thorley, 1984; David, 1987; Rowe, 1988; Wilson and Scott, 1989; Pollock and Warner, 1989; Sarantinos et al, 1990; Rowe and Rowe, 1994; Batemann et al, 2004; McCann et al, 2007). This group of clinical trials assessed the validity of Feingold's hypothesis relating artificial colors and ADHD and other problem behaviors in susceptible children. The first fourteen of these trials (Conners et al, 1976 to Spring et al, 1981 as listed above; also color-coded grey† in Table 1) were conducted prior to 1982, at which time the National Institutes of Health convened a Consensus Panel to review the status of information regarding defined diets (Feingold diet) and childhood hyperactivity (NIH, 1982). Subsequent to that NIH Consensus Panel, ten additional clinical trials (Salamy et al, 1982 to McCann et al, 2007, as listed above) were conducted to investigate further the adverse behavioral effects of artificial colors in children. Group II consists of 7 clinical trials (color-coded blue† in Table 1) that are characterized by the use of an experimental treatment consisting of either a foods elimination diet (also referred to as 'oligoantigenic' and 'few foods' diet that excludes all foods, additives and food components assumed to provoke adverse behavioral reactions in certain children) or challenges with specific provoking food items (Egger et al, 1985; Kaplan et al, 1989; Egger et al, 1992; Carter et al, 1993; Boris and Mandell, 1994; Schmidt et al, 1997; Uhlig et al, 1997). The use of this type of experimental treatment in studies, all of which were conducted after the NIH Consensus Panel (NIH, 1982), focused on a broader array of substances by assessing adverse effects of food itself in hyperactive and problem behavior children and reflected the growing opinion that it may be more likely that any adverse food related behavioral changes in hyperactive children are caused by an individual intolerance to various foods or food components (which could include color additives), rather than by a singular involvement of color or other additives (Arnold, 1999; Bishop, 1983; Courmier and Elder, 2007; Robinson and Ferguson, 1992; Stare et al,

1980;Young, 1997). The two different experimental treatments in these two groups of clinical trials may reflect either two different views or one evolving view of the relationship between food/colors and ADHD and other problem behaviors in children.

All of the clinical trials were designed to be conducted under double-blind, placebo-controlled conditions using one of two basic types of experimental procedures - "*diet crossover*" which focused on determining under controlled scientific conditions whether a particular defined diet was an effective dietary treatment for reducing problem behaviors in children, particularly those associated with ADHDs, and "*specific challenge*" which focused on determining whether exposure to a particular food item, excluded from the defined diet, could trigger or exaggerate adverse behavioral changes in the test population of children (Bishop, 1983).

Group I Trials (n=26)

As noted above, the focus of these trials was to assess the validity of Feingold's hypothesis relating artificial colors and ADHD and other problem behaviors in susceptible children. Of the 26 trials in this group, 2 were "*diet crossover trials*" assessing the effectiveness of the Feingold diet, eliminating artificial food colors, flavors and natural salicylates, in improving the behavior of hyperactive children. The remaining 24 trials were "*specific challenge trials*" which assessed the adverse behavioral effects of artificial food colors (either mixtures or single colors), occasionally in conjunction with a food preservative, in children diagnosed ADHD, with problem behaviors, or from the general population. Although Feingold's hypothesis implicated artificial colors, flavors, natural salicylates, and other additives, clinical investigations focused initial attention primarily on artificial food colors, since colors were specifically mentioned by Feingold as being of primary concern and since colors represented a relatively smaller group of chemicals to deal with in comparison to the food flavors and were thought to be more easily masked (blinded) than artificial flavors in the challenge studies. The implied intent, however, seemed to be that eventually other categories of food additives would be investigated. With the exception of several artificial color studies that also included one or two preservatives as part of the experimental treatment, no systematic clinical investigation of adverse behavioral effects of flavors or natural salicylates has yet been carried out.

Typically, the procedure for the Group I "*diet crossover trial*" involved randomly assigning a group of children to either a Feingold diet or a matched placebo diet for a defined period of time, then crossing them over to the other diet condition for an equivalent period of time and evaluating the behavioral responses to both diet conditions. For the "*specific challenge trial*" all subjects were maintained on the Feingold diet for a variable period of time prior to and throughout the study period. Each child was given a color challenge (i.e., a blend of colors or an individual color) in some delivery item that masked the presence of the color (e.g. cookies, cupcake, drink, or capsule), and a matched placebo treatment in randomized fashion for a defined period of time or on random days. A variable washout period was used between challenge periods. Behavioral responses were measured at specified times for both treatment conditions. It should be noted that challenge studies provide information for only one component part of the Feingold diet (e.g. artificial colors) and should not be interpreted as evidence of effectiveness for the diet as a whole. In the recruitment of subjects for the "specific challenge" trials, many trials (18/22 cited below) included efforts to maximize detection of behavioral

effects in the controlled challenge study by including a specific selection criterion that the children had to be reportedly sensitive to the Feingold diet under non-blind conditions, showing marked behavioral improvements on the open Feingold diet and exhibiting immediate and dramatic deterioration of behavior after ingesting even small amounts of prohibited foods containing artificial food colors and other additives (Harley et al, 1978- Phase II; Goyette et al, 1978; Rose, 1978; Levy and Hobbes, 1978; Levy et al, 1978; Mattes and Gittelman-Klein, 1978; Conners, 1980; Conners et al, 1980; Weiss et al, 1980; Mattes and Gittelman-Klein, 1981; Adams, 1981; Spring et al, 1981; Salamy et al, 1982; David, 1987; Rowe, 1988; Wilson and Scott, 1989; Pollock and Warner, 1989; Rowe and Rowe, 1994)

Individual Trial Summary Reviews (Group I)

- In the first controlled crossover trial of the Feingold diet **(Conners et al, 1976)** the teacher behavior ratings showed an overall behavioral improvement for a group of 15 school age diagnosed hyperkinetic children on the Feingold diet but the biological significance of this finding was confounded by an inexplicable treatment x order effect in which the improved behavior was seen only when the placebo diet was given first. Also, the parent ratings showed no diet related changes in behavior. Viewing the data for individual subjects indicated that teacher and parent ratings did agree that 4 to 5 of the 15 subjects showed improved behavior on the Feingold diet. Since the significance of even this observation was also confounded by the treatment x order effect, overall this trial provided no evidence that Feingold's diet improves behavior of hyperkinetic children.

- The only other controlled crossover trial of the Feingold diet was conducted by **Harley et al (1978: Phase I)** who included independent classroom observations and objective laboratory testing in addition to parent and teacher behavior ratings to assess behavioral change. This trial tested 46 hyperactive children (36 school age and 10 pre-school age). Analysis of the data for the 36 school age children showed inconsistent findings, similar to Conners et al (1976), except that here the behavioral improvements on the Feingold diet were noted by the parent ratings but were not confirmed by either the teacher ratings or the other outcome measures used in this study. The biological significance of the parent finding was also confounded by an inexplicable treatment x order effect. This also confounded the observation that parents and teachers agreed on a diet response in 4 of the 36 children. Analysis of the data for the 10 pre-school children also showed inconsistent findings with significant behavioral improvement on the Feingold diet being noted by the parent ratings but no confirmatory diet effect being found with either the classroom observation or laboratory testing (teacher ratings were not available for the pre-school children). In contrast to the school age children, the parent rated diet change in behavior of the pre-school children was not confounded by a treatment x order effect. Overall, in view of the inconsistent findings and the confounding treatment x order effect, this study provides no credible support for the contention that the Feingold diet affects disruptive behaviors in either school age or pre-school hyperactive children. It is questionable whether these findings would support the suggestion that a small subpopulation of younger, pre-school age hyperactive children may be responsive to the Feingold diet.

- In one of the first uses of the specific challenge design to assess the adverse behavioral effects of artificial colors, **Goyette et al (1978)** conducted a two-part trial. In Trial 1, 16 hyperactive children (4.7-11.8 years, mean age 8.3), prescreened to be 'diet responsive'

(behavioral improvement on open, i.e., non-blind, Feingold diet), exhibited no significant behavioral response to a daily cookie challenge with 27 mg mixed colors for two 2-week periods based on multiple outcome measures (parent rating, teacher rating, laboratory tracking test). A non-significant trend of performance deficit related to color challenge was noted for the younger subjects at 1 but not 2 or 3 hours after treatment. Trial 2 was a repeat of Trial 1 except that: a parent rating focused on the 3 hours after treatment was the only outcome measure; daily challenge was 13 mg mixed color for 1 week; and 13 younger children (3.4-10.2 years, mean age 6) were recruited, 8 "criterion" diet responsive/hyperactive (3.4-8.4 years, mean age 5.3) and 5 "borderline" diet responsive/hyperactive (3.7-10.2 years, mean age 7). Significant color challenge effects were found based on the parent ratings and, notably, without any differences in color responses between the 'criterion' and 'borderline' groups. In the absence of any other outcome measure to confirm the parent finding and without any verification of effectiveness of blinding, the reliability of the Trial 2 findings are considered questionable. Also, there seems to be limited support for the conclusion that artificial colors may be particularly disruptive to younger children or that they are more sensitive to artificial colors. The suggestive trend in Trial 1 of a treatment effect on tracking performance in younger subjects is not significant (and not supported by the parent and teacher ratings) and in Trial 2, the limitations of these findings notwithstanding, the 'criterion' children (mean 5.3 years) were younger than the 'borderline' group (mean age 7 years), yet there was no difference between these two groups in their color responses. Overall, this study does not provide credible evidence of an effect of artificial colors on behaviors of reported diet responsive hyperactive children. Note that a replication of this study by Conners (1980) detected no significant adverse effects of daily challenge with 26 mg mixed colors and the investigator concluded that the effects reported by Goyette et al (1978) were most likely due to chance and not biologically relevant.

- **Harley et al (1978 :Phase II)** conducted color challenge trial, as the second phase of an earlier diet crossover trial, using the best 9 diet responding hyperactive subjects from that earlier study. There was some question whether these children still appeared to demonstrate appreciable behavioral improvement when placed on the open Feingold diet immediately prior to the challenge phase of this study. A daily challenge with 27 mg mixed color in candy bars/cookies over 2 periods of 2 or 3 weeks each produced no significant behavioral changes based on multiple outcome measures (parent, teacher and trained classroom observer ratings and neuropsychological testing). There was a tenuous, but suggestive observation, that the parent rating and observer rating profile 1 of the 9 subjects may have approximated that of a color challenge effect, but this was not supported by the teacher ratings (individual neuropsychological data were not inspected). This study does not support the hypothesis that artificial food colors cause or trigger hyperactivity behaviors of childhood.

- **Levy and Hobbes (1978)** conducted one of the first color challenge studies using a single artificial color, tartrazine, as the treatment. Parent ratings and global assessments revealed no significant adverse behavioral effects of daily challenge with 4 mg *tartrazine* in cookies for 2 weeks on behaviors of 7 hyperactive children, identified as responsive to the Feingold/KP elimination diet. Within the context of the limitations for this study, that no blinding procedures were described and only one source (parent) of outcome determination was used, this study provides no evidence that tartrazine has any adverse effect on behavior in hyperactive children who are reportedly responsive to the Feingold diet.

- **Levy et al (1978)** conduct a challenge study of tartrazine (Yellow 5) in 20 hyperkinetic children. Responsiveness to the Feingold diet was determined in open trial prior to the challenge study. In group analysis there was no significant adverse behavioral effect of daily challenge with 5 mg tartrazine in biscuits for 2 weeks based on multiple outcome measures (mother, teacher and clinician ratings, and a series of objective psychological tests of activity, attention, impulsivity, perceptual-motor function, memory, intelligence and maze performance). Additional post-hoc analyses determined that the level of hyperactivity (prior to study) did not amplify the children's sensitivity to diet (i.e., response to open treatment with Feingold diet) or to tartrazine challenge. Additional post hoc analysis was conducted on a subgroup of children specifically rated by mothers under non-blind conditions as diet-responders, i.e. exhibited significant reductions in problem behaviors when placed openly on the Feingold diet. Analysis of the challenge data for these children did show a significant adverse tartrazine challenge effect based on mother ratings. However, no other confirmatory outcome measure was included in this particular post-hoc analysis. The blinding of this study may have been compromised since some of the color challenge biscuits differed in color from the placebo biscuits. Considering that the main analyses of the study data showed no significant color challenge effect based on multiple outcome measures, that the blinding in this study was compromised to some indeterminable extent, that the singular treatment effect was based on post hoc analyses of only one source of outcome measure (mother ratings) with no other outcome measures analyzed to support this finding, there is, overall, little, if any, confidence in the reliability or biological relevance of this study's finding of a treatment related effect or in its value to suggest that artificial colors have any effect on hyperactivity behaviors in childhood.

- The controlled challenge trial reported by **Mattes and Gittelman (1978)** using a mix of common artificial colors is significant for two reasons. This was the first attempt to individualize dosage of the challenge material to the sensitivity of the subject and this may have been one of the earliest suggestions that adverse effects of artificial colors may be associated with general behavior disturbances such as irritability and restlessness/fidgetiness, which are not typically characteristic of childhood hyperactivity. The single subject used in this trial was diagnosed hyperkinetic who reportedly showed improved behavior on the open Feingold diet and dramatic deterioration of behavior after eating prohibited food, with restlessness and irritability lasting hours to days. The initial "dose-ranging" phase of this trial was conducted under apparently blinded conditions to determine an appropriate behaviorally effective daily dosage level of the color mix to be used in the subsequent challenge phase. Although parents and teachers rated the child's behavior during the "dose-ranging", only the parent ratings were able to detect any color related change in behavior (the mother's Conner rating score went from a baseline of 0 to15, the cutoff index for hyperactivity). During the subsequent 10 week double-blind multiple crossover challenge phase, the child was given cookies with approximately 18 mg mixed colors on two days per week randomized with placebo weeks. Neither parent nor teacher ratings were able to detect any adverse effects of color challenge on behavior. However, the mother was able to guess correctly which treatment was given for 8/10 weeks, focusing on irritability and fidgetiness as the main changes. In view of the inconsistencies between the mother's ratings during dose-ranging and challenge testing, her ability to "guess" correctly most of the weekly treatments, and the fact that the teacher detected no color related change in behavior at anytime during dose-ranging

or challenge testing, the findings from this study are not readily interpretable and provide no credible support of an association between artificial colors and hyperactivity behaviors in children.

- **Rose (1978)** conducted a challenge study of the artificial color, tartrazine (Yellow 5), using a within-subject design with a sample of 2 hyperactive children reported by parents to be behaviorally responsive to the Feingold diet and maintained on that diet for behavioral management. Over a period of 5 weeks, children were given one oatmeal cookie daily. On one day during each of two different weeks (weeks 2 and 4) children were given the color challenge of a single cookie with 1.2 mg tartrazine. The sole analyzed outcome measure, provided by a trained observer, was an assessment in the school environment of specific behaviors characteristic of the hyperactive syndrome. While considerable efforts were made to ensure the blinding conditions, no actual verification of the effectiveness of the blind for the observer, parents and children was carried out. The analyses showed that for each subject there was a significant increase of hyperactive behaviors associated with the tartrazine challenge. Without the availability of other outcome measures to confirm or support these findings and some uncertainty about the effectiveness of the blinding, the study is viewed as suggestive of a functional relationship between tartrazine and hyperactive behavior in children reportedly responsive to the Feingold diet. Although with this limited sample of 2 children it is difficult to extrapolate the significance of these findings.

- **Williams et al (1978)** designed a rather complex challenge study to focus on the relative effects of a diet free of artificial food colors, flavors and major preservatives (i.e., a modified Feingold diet) compared to stimulant medication (methylphenidate) in managing hyperactivity in children. As a point of clarification, however, this study did not specifically assess the effects of "diet" on behavior. Rather, the study conducted was a double-blind challenge trial assessing the behavioral effects of stimulant medication and a mixture of artificial food colors, administered separately and in combination, in a heterogeneous group of 24 hyperactive/problem behavior children who were all maintained on a modified Feingold diet throughout the study. Prior reported responsiveness to the modified Feingold diet was not a criterion in subject selection for this study, but responsiveness to stimulation medication was a criterion. Over a period of 4 weeks each child was challenged on 4 days each week with a different combination of stimulant medication and mixed colors (26 mg in chocolate cookies) or their respective placebo treatments. Based on analysis of the parent and teacher ratings (standard Conners questionnaires), there were clearly significant effects (behavioral improvement) for the stimulant medication challenges but the effects for the artificial color challenges were inconsistent. Teacher ratings revealed a significant adverse behavioral effect of color challenge, but this was not confirmed by the parent ratings which found no significant effects. In the post hoc analysis of the rating data for individuals, the parent ratings identified 3 "criterion responders", i.e. subjects showing a 33% or greater behavioral deterioration with color challenge compared to placebo, and the teacher ratings identified 5 "criterion responders". But, these findings were inconclusive since there was no agreement between the parent and teacher lists of responders. Overall, in view of the clear effect noted by both parents and teachers of stimulant medication in improving hyperactive/problem behaviors, this reviewer views the disparity between the parent and teacher findings, in terms of the main findings of color challenge effects and the identification of different color responsive children, as definitely problematic for

interpretation of the color challenge findings and considers the study to be somewhat suggestive but inconclusive and not supportive of an effect of artificial colors on hyperactive and problem behaviors in children.

- **Conners (1980a)** conducted a challenge study of the behavioral effects of artificial colors to replicate a prior study by Goyette et al (1978) but with a larger group of 30 diet responsive hyperkinetic children. Based on parent and teacher ratings, there were no significant adverse behavioral effects of challenge with 26 mg mixed colors administered daily for two alternating 1-week periods in hyperkinetic diet responsive children. Since these negative effects produced no confirmation of the color challenge effects reported in Goyette et al (1978), this reviewer agrees with the conclusion that the effects reported by Goyette et al (1978) were most likely due to chance and not biologically relevant.

- **Conners et al (1980)** attempted to determine whether hyperactive children who appear to be sensitive to artificial food colors can be shown to exhibit a time dependent response to those colors using sensitive laboratory tests and observations. The 9 hyperactive children selected for the double-blind challenge trial were reported to exhibit under non-blind conditions marked behavioral improvement on the Feingold diet and notable behavioral sensitivity to foods with artificial colors. While maintained on the Feingold diet, each child was randomly given 2 challenge sessions at 1-2 week intervals, one session using color cookies (30 mg mixed colors in chocolate cookies) and the other session placebo cookies. Each session consisted of baseline measurements, followed by the appropriate treatment and then retested at 45, 90, 135 and 180 min after challenge. Testing consisted of behavior ratings by the experimenter and objective tests including actometer readings, activity chair measures, and a paired-associate attention/learning task. No significant adverse effects of the color challenge were found with any of the test measures. The investigators suggested several factors that may have minimized or confounded the detection of any significant color challenge effects, including the low detection power of the test with small numbers of subjects, the use of an ineffectively low level of artificial colors (30 mg), the increased levels of activity with ingestion of both color and placebo cookies masking any specific effect of the color, and particularly the marked practice effect for the paired-associated task across sessions. The suggestion that these factors may explain why treatment effects were not detected seem inconsistent with the fact that all of the children in this study were selected specifically based on teacher and/or parent reports, although non-blinded, of marked behavioral changes with the Feingold diet and notable sensitivity to artificial colors, and not merely subtle laboratory changes. This selection criterion was intended to enhance the ability to detect a real treatment related effect if there was such an effect. Within this context, it seems clear that this study failed to demonstrate any significant effects with in-depth testing at specified intervals up to 3 hours after challenge with a 30 mg mixture of artificial food colors in hyperkinetic children who were reportedly diet responsive and color sensitive.

- Attributing the poor response in previous studies to insufficiently low levels of colors, **Swanson and Kinsbourne (1980)** designed a challenge study to assess the effects of larger challenge dosages of artificial food colors in a hospital setting. This trial involved 20 children diagnosed hyperactive (average Conners score 16.2) who were responsive to stimulant medication (*hyperactive set*) and 20 children with problem behaviors not rated hyperactive (average Conners score 12.3) and who were reported to have adverse effects to stimulant

medication (*non-hyperactive set*); neither set was identified as diet responsive but all children were maintained on the Feingold diet prior to and during the challenge period. On one of two test days, ten children from each *set* were given a single capsule with 100 mg blended colors and the remaining ten children from each *set* were given a capsule with 150 mg blended colors; a placebo sugar capsule was given on the other day. The outcome measures included paired associate learning testing (PAT) which was assessed three times within 3.5 hours of treatment and Conners behavior ratings assessed two times each test day. Significant color related decrements in PAT performance were found following both dosage levels of 100 and 150 mg but only for the *hyperactive set* of children and not for the *non-hyperactive set.* There was no difference in effect between the two treatment levels of color, indicating no dose-response at least at these levels. This apparent effect of color challenge on performance of PAT was not supported by the Conners behavioral ratings which were described as (no data presented) showing no differences between the color challenges and placebo treatment. However, the reported color effect on PAT performance in the hyperactive children may be questionable due to an inexplicably unusual placebo performance specifically of the hyperactive children (the placebo performance of the hyperactive children improved across test intervals but the placebo performance of the non-hyperactive children deteriorated across intervals). Confidence in the findings of a color challenge effect in the pair-associate test may be further limited by an uncertainty about the effectiveness of the blind in this study, since the use of simple sugar capsules as placebo may not adequately blind for the presence of 100/150 mg color and no procedure was described to verify that the study personnel or children were effectively blinded to treatment. Also, the Conners ratings provided no supportive evidence of any behavioral effects of the color challenge. The investigator's suggestion that the objective PAT performance is more sensitive than the subjective Connors behavior ratings for documenting adverse effects of color on behavior seems inconsistent with the fact that the investigators considered the Conners score sensitive enough to identify the study populations of subjects in this study as *hyperactive* or *non-hyperactive* and that other investigators have found significant treatment effects with parent/teacher behavior ratings but not with objective testing (Harley et al, 1978; Egger et al, 1985; Carter et al, 1993; Batemann et al, 2004). Overall, in view of the questionable effectiveness of the blinding in this study, the absence of supportive evidence of any color related changes based on the Connors ratings, and the inexplicable difference in placebo performances of the paired-associate test by hyperactive and non-hyperactive children, there is minimal confidence in the reliability of the reported findings and this reviewer considers this study to be inconclusive and to provide no credible or suggestive support for a proposed association between hyperactivity or other problem behaviors and artificial food colors in children.

- In consideration of the possibility that behavioral responses to artificial colors may be specific to individual children, the challenge study conducted by **Weiss et al (1980)** was the among the first to include parental rating of selected behaviors targeted for each child as one of the primary outcome measures of treatment effect. The 22 subjects in this trial were non-hyperactive children with behavioral problems, whom parents reported to be responsive to open (non-blinded) treatment with a modified Feingold diet on which the children were maintained throughout this study. On 8 randomly selected days over a 77 day study period, children were challenged with a soft drink containing a 35 mg blend of artificial colors. Matched placebo drinks were given on all other days. Behavioral assessments were made by the parents' targeted ratings, Conners ratings, continuous records of aversive behaviors, and global estimates of target behaviors. Data for each child was analyzed separately (i.e.,

considered individual experiments). Under conditions designed to maximize detection of adverse effects of artificial colors, i.e., use of assessments including individualized targeted behavior ratings and use of subjects with problem behaviors who reportedly improved on a modified Feingold diet, a dramatic adverse response to color challenge was found in one of the 22 children tested. The remaining 21 children showed no convincing adverse reaction to the 35 mg color challenges, including one child showing only a very limited and inconsistent response. Any conclusions from these results should reflect the facts that the detection of treatment effects was based on only one source of behavioral assessment, the parents' observations, and that there was no specific validation of the effectiveness of the blind for the parents or children (e.g., asking them if they could differentiate the two treatment items based on some physical characteristic). In the absence of any other source of behavioral assessment to confirm or support the parents' assessments, confidence in the significance of the study findings is limited. While the findings do not provide definitive evidence, these findings may be viewed as suggesting that a small subgroup of children with problem behaviors who reportedly improve on a modified Feingold diet may be intolerant to artificial color additives, exhibited by adverse behavioral effects. Since all of the children in this study were reported to have improved on a modified Feingold diet, although in non-blinded conditions, and most of these children showed no sensitivity to artificial color additives, this would suggest that intolerances to various foods or food ingredients, other than colors, are of more significance in provoking or aggravating problem behaviors. This in fact is reflected in the investigators' conclusion that "…. colors, and perhaps other agents excluded by elimination diets, can provoke disturbed behavior in children".

- **Mattes and Gittelman (1981)** incorporated several elements in the design of this crossover challenge study to maximize the likelihood of detecting any adverse behavioral effects of the color challenges in a group of children heterogeneous for hyperactivity (6 ADHD/ADD and 5 problem behaviors or history of hyperactivity). Specifically, (1) this study included only children reported by non-blinded parents to show marked behavioral improvement on the Feingold diet and to deteriorate quickly and dramatically with exposure to foods with artificial colors, (2) made efforts to exclude children responding to the placebo cookie alone, and (3) used high dosages of artificial colors as challenge (levels of the blended colors increased daily during the one challenge week from 13 mg/day to 78 mg/day). Based on an extensive battery of subjective (parent, teacher and psychologist behavior ratings, and psychiatric evaluation) and objective (laboratory test of distractibility) behavioral measures, there were no overall significant differences between placebo and the incremental color challenge. Viewing the data for individual subjects, parent ratings did indicate that six children did show some color challenge effects. However, this was not confirmed or supported by any of the other behavioral ratings/evaluations (teachers, psychologist, psychiatrist) all of which showed no treatment related differences in the behavior of any of the children, making the clinical relevance of the parent ratings highly questionable. The results of this study indicate that artificial food colors, even at relatively high exposure levels, have limited, if any, adverse effects on the behavior of most school-age children with hyperactivity or other behavioral problems who are claimed by parents to be markedly sensitive to these agents. Assuming some level of credibility to the non-blind reports of behavioral improvement of the children in this study when they were placed the Feingold and of dramatic behavioral deterioration with exposure to foods with artificial food colors, the basically negative results from this study might suggest that some foods or food ingredients,

other than artificial colors, may be more significantly associated with provoking or aggravating problem behaviors in some hyperactive or problem behavior children.

- **Adams (1981)** conducted a double-blind, placebo-controlled challenge trial to investigate whether an *infraction* of the Feingold diet under controlled clinical conditions exaggerates hyperactive behaviors in children. The trial included a total of 18 hyperactive children who were reported to be responsive to the Feingold diet and sensitive to infractions of the diet. The challenge *infraction* was an iced chocolate cupcake with a glass of lemonade both made from commercial mixes with added artificial red/yellow colors (and vanilla flavor) to produce an "artificial snack" containing 26.3 mg mixed food colors in addition to other unidentified additives; the placebo ("natural snack") cupcake and lemonade were made with color-free natural ingredients. Parents and observer was blinded to which snack was given and, while the effectiveness of this blind was not verified, there was no indication that integrity of outcome measures was affected. While being maintained on the Feingold diet, each child was randomly exposed to both the "artificial snack" and the "natural snack", one each on two separate days (interval between days was not specified). On each test day, the children ate their assigned snack 3 to 4 hours before being tested with a battery of objective behavioral tests of activity, memory, attention and motor skills. As a second source of behavioral assessment, on each test day the parents also rated the degree of change in their child's behavior during the 3 to 4 hour period after the snack was consumed (no specific parental questionnaire was described). Data for individual subjects were not analyzed. The results of the group data analyses showed no significant main or interactive effects of the "artificial snack" *infraction* on any of the battery of objective behavioral tests or on the parental rating of degree of change in behavior. Although data for individual subjects were not analyzed statistically, the investigator did mention that only small or slight changes in behavior were noted for any given child. Although the effectiveness of the blinding for the parents and children was not verified in this study, there was no reason to believe that the blinding may have been compromised or otherwise affected the credibility of this trial's results. Overall, the primary finding from this trial was that, despite the very select nature of the hyperactive diet-responsive subjects, no significant overall adverse behavioral effects resulted in these children from a single *infraction* of the Feingold diet, based on a battery of objective tests and subjective parental ratings. Since the only artificial ingredients identified in the food items used as the dietary *infraction* included 26.3 mg mixed artificial red/yellow colors and an unknown amount of artificial vanilla flavor, it may be concluded only that this study's findings provide no evidence of an association between ingestion of these particular artificial ingredients and behaviors associated with hyperactivity in a select group of school-aged hyperactive children. While this study does not disprove the possibility that there may be behavioral effects produced by some food substances in some susceptible children, it may lend support to the view that, if diet is related in some way to problem behaviors, this effect is less reliably demonstrated than parents predict and may possibly call into question the reliability of a parent's belief of what substances negatively affect their children's behavior.

- **Spring et al (1981)** conducted a controlled challenge trial to test the Feingold hypothesis, recruiting only hyperactive children (n=6) already on the Feingold diet whose mothers reported a diet-related reduction of hyperactivity symptoms. Additionally, to ensure that any treatment related findings were causal and not coincidental effects, a replication procedure similar to the original experimental procedure was also carried out. The active challenge item

was a chocolate cookie containing a 13 mg blend of eight approved artificial food colors; subjects received two cookies per test day, resulting in a total daily color challenge of 26 mg. Placebo cookies contained no artificial colors. Independent testing confirmed blinding of challenge cookies. With the Feingold diet continued, children were given two cookies daily, one before school and one after school, on Tuesday, Wednesday and Thursday during each of six weeks (primary experimental period). Active challenge cookies were given during two of those six weeks. Three randomly selected children received active challenged cookies during weeks 1 and 4; the remaining three children received active challenge cookies during weeks 3 and 6. Placebo cookies were given during the other weeks. On each cookie day, both the mother and teacher were asked to rate their child's behavior using two outcome measures. One measure was an abbreviated 10-item modified version of the Conners hyperactivity rating scale, which the investigators verified as a valid measure of hyperactive behavior but for which the score level that indicated hyperactivity was unclear. For the second outcome measure (global behavior judgment) the respondents (mother and teacher) were asked to guess which treatment the child received based on their assessment of the child's general behavior on that day. Each of the 6 subjects constituted a single subject experiment and analysis of the grouped data was not conducted. Subsequent to the original trial, the investigators conducted a replication procedure that was similar to the original trial but collected only the mother's rating data (hyperactivity rating and global behavior judgment). When the data from the primary experiment were analyzed with the two active challenge periods combined, the results indicated that only two of the six subjects appeared to respond to the active challenge. For one of these two subjects (Subject A), the only indication of a treatment effect was the mother's hyperactivity ratings which were significantly correlated with active challenge but which were not confirmed by the teacher's ratings. Subsequently, it was revealed that a home event occurred that may have influenced the mother's ratings. Consequently, the investigators essentially dismissed the findings for Subject A since they considered it likely that the relationship between the mother's rating of increased adverse behaviors for Subject A and the active color challenge may have been coincidental rather than causal. For the second apparent responding subject (Subject E), significantly accurate guesses were made only by teacher but were not confirmed by the mother. However, the daily hyperactivity ratings for Subject E showed that both the mother's and teacher's ratings were significantly correlated with the active challenge. The relevance of these latter correlations became suspect when the hyperactivity ratings data for Subject E were viewed separately for each week of the study period, since the ratings by both the mother and the teacher appeared (to this reviewer) to be inconsistent between challenge periods and the raters did not appear to agree on which challenge period the change in behavior occurred. The absence of any significant effect of the active color challenge on the behavior of Subject E was confirmed in the replication procedure, based on the mother's daily hyperactivity ratings and global behavior judgment (daily guesses about treatment). In assessing the overall findings of this study it should also be noted that, although the effectiveness of blinding was not verified for parents, teachers or children, there was no indication that the credibility of the outcome measures was in any way affected. Overall, this reviewer's assessment concludes that this study's findings provide no consistent or credible evidence that hyperactive behavioral effects were elicited or exaggerated by repeated challenge with 26 mg mixed artificial food colors during two separate weeks in a small sample of hyperactive boys. It is possible that these children, selected for their reported sensitivity to the Feingold diet, could have been responsive to food elements in that diet other than artificial colors.

- **Salamy et al (1982)** designed a single challenge clinical trial using physiological measures of autonomic nervous system (heart rate) and central nervous system (electroencephalogram) activity for evaluating the effects of the artificial food color, Red 40. Four hyperactive children who were reported to be responsive to the Feingold diet and four non-hyperactive sibling controls were maintained on the Feingold diet during this study. Over two days children were randomly treated with a commercial drink colored with Red 40 (level of Red 40 or other additives in the drink were not specified) and a matched placebo. There was no evidence of significant physiologic changes (HR or EEG) attributable to the artificial color/additives treatment in either the hyperactive or control subjects. But there was a striking difference between the two groups of subjects, with a relative absence of physiological changes in the control subjects but an abundance of changes in the hyperactive children during both testing sessions unrelated to treatment. The hyperactive children seemed to be more labile physiologically, responding generally to the experimental situation. Even though these results were negative, the absence of information regarding the level of Red 40 used for the challenge does not allow the possibility to be dismissed that the level used may have been too low to elicit physiologic responses particularly with the few numbers of subjects tested.

- **Thorley (1984)** conducted a double-blind clinical challenge study designed to maximize detection of any behavioral and cognitive response to color challenge, including such features as a high challenge dosage level of 92 mg of blended colors, optimized testing using a battery of sensitive subjective and objective outcome measures, use of a residential setting for better control, and a relatively homogenous population of intellectually retarded children with inattentive/hyperactive behavior. The 10 subjects, none of whom were identified as diet responsive, were maintained on a modified Feingold diet and given a cocoa drink daily throughout the 28 day study period. Each child was randomly selected and given the color challenge (92 mg blended color in a cocoa drink) on 2 consecutive days. Outcome measures included teacher ratings which included a devised individual rating scale of most problematic behaviors, care staff ratings, psychometric testing with mazes and paired associate learning test, and actometer recordings. No significant color related effects were found for any outcome measure. even though there may have been some question regarding the effectiveness of blinding and the appropriateness of analyzing the data using uncorrected repeating t-testing, the absence of any adverse effect of a high dosage (92 mg) of artificial colors in a study designed to maximize detection of behavioral and cognitive responses indicates that it would be unlikely that intellectually retarded children as a whole would show adverse effects to artificial food colors.

- **David (1987)** conducted a double-blind challenge study using high dosages of tartrazine (50 mg and 250 mg) and benzoic acid (50 mg and 250 mg) to verify parent reports that their children showed immediate adverse behavioral reactions (within two hours) to foods with additives, particularly tartrazine and benzoic acid. The 24 children in this study were heterogeneous with problem behaviors related to hyperactivity and were maintained on elimination diets that avoided food additives and any other suspect foods. Twelve were inpatients in a general pediatric ward and 12 came to the ward as outpatients. Each child was challenged with a drink containing 50 mg tartrazine and two hours later challenged with a second challenge drink containing 250 mg tartrazine. Subsequently, but on a separate day, all subjects were challenged with drinks containing the same two dosages of benzoic acid

following the same procedure. Apparently, no randomized treatment order was used, either between tartrazine and benzoic acid challenge periods or between each of their dosages. The outcome measures consisted of recorded observations by a parent and the nursing staff, but no specific or structured scoring system was used and it was not specified how or whether data were statistically analyzed. Parents of outpatient children knew when challenge days occurred but parents of inpatient children and nursing staff did not. Based on recorded observations by parents and nursing staff, none of the children was reported to show any treatment related change in behavior following the tartrazine or benzoic acid challenges. Unfortunately, there were several rather prominent limitations and weaknesses in the design and conduct of this study. These included the facts that the outcome measure consisted of unstructured observations, randomized treatment conditions were not used, half of the parents were not blind to active challenge days, and data were neither presented nor statistically analyzed. While these issues raise questions about the sensitivity and reliability of this study for detecting little more than obvious treatment related effects, this study was able to show that there was no confirmation of the parents' anecdotal reports of immediate and obvious adverse behavioral reactions to foods specifically attributable to either tartrazine or benzoic acid. It may be possible that those parental reports, although anecdotal, of a history of obvious adverse behavioral changes in these children may have been associated with some food items or components other than tartrazine and benzoic acid.

- **Rowe (1988)** conducted a controlled challenge clinical trial to assess the behavioral effects of two artificial colors, tartrazine and carmoisine, in 8 children who were suspected hyperactive and who reportedly showed improved behavior in an open trial (Phase I of this study) of the Feingold diet and reacted adversely (commonly described as extreme irritability, restlessness and sleep disturbance) to the ingestion of foods containing additives, particularly the red and yellow artificial food colors. All children were maintained on Feingold diet and received one capsule daily during the 18 week study. Each child, in a random sequence, was given daily capsules with 50 mg tartrazine or 50 mg carmoisine each for one week on 2 randomly determined occasions (i.e., a total exposure of 2 weeks for each color) with 2 or 3 weeks washout between each color week. Parents and children were blind to challenge sequence but verification of effectiveness of blind was not determined. Data for each child were analyzed separately. Based on the sole outcome measure of daily parent behavior ratings, using a targeted frequent symptoms checklist for each child, only two (one hyperactive, the other non-hyperactive with problem behaviors) of the 8 subjects were identified as showing significant adverse behavioral responses to both the 50 mg tartrazine and 50 mg carmoisine challenges. Their responses (increased activity, irritability, low frustration tolerance, sleep disturbance, short attention span) began within hours after initiating each color challenge treatment and continued for several days to several weeks after the last of each color challenge treatment week. Both children were atopic, but symptoms did not relate to color challenges. Attempted teacher ratings were incomplete. Confidence in the reliability of these positive findings is limited by the absence of any other source of behavioral measures to corroborate the parental ratings and the lack of attempt to verify the effectiveness of the blind for parents and children. While not conclusive, these findings do suggest that a small subgroup of children with problem behaviors heterogeneous for hyperactivity may be adversely affected by the artificial colors, tartrazine and carmoisine, and that this subgroup of color intolerant children may not be limited to those with ADHD. Interestingly, the common behaviors parents emphasized as being associated with open ingestion of foods with colors, and possibly other intolerant food substances, were extreme "irritability", "restlessness" and

"sleep disturbance" rather than those associated with "attention deficits" (suggested by Mattes and Gittelman, 1978). The author viewed this as suggesting that the inclusion of children in clinical trials on the basis of 'attention deficit' alone may miss some color reactors. The six children who did not react to the controlled color challenges were reported in open (non-blind) trial to be diet responsive and to react adversely to foods containing food additives. It is possible that part of these apparent responses to diet may have been associated with intolerance to certain food items/components eliminated from the Feingold diet other than artificial colors. This suggests a broader view of the possible behavioral impact of artificial food colors, and possibly other food intolerances, from only children with ADHD and related behaviors to children with more general types of problem behaviors, including but not limited to those associated with ADHD.

- **Wilson and Scott (1989)** conducted a controlled clinical challenge trial to assess food additive intolerance in 19 non-hyperactive children reported to be responsive to open (non-blind) treatment with a color-free diet and with a definite but anecdotal history of adverse reaction (respiratory, dermatologic, and/or behavioral) to artificial yellow color. For most children color intolerance was verified in open (non-blind) challenge prior to the study. All subjects were maintained on an additive-free diet throughout the study. Two types of challenge drinks were use: a color challenge aide drink with 8.5 mg tartrazine and 8.5 mg sunset yellow, and a preservative challenge aide drink with 12.5 mg sodium metabisulphite and 55 mg sodium benzoate. The placebo treatment was an aide drink with B-carotene coloring to match the challenge treatments. The initial challenge stage of the double-blind trial consisted of three 14-day phases with each of the three experimental treatments randomly assigned double-blind to the three phases for each child. During each phase one bottle of the same type of drink was consumed daily for 12 days followed by a 2-day washout period. If behavioral symptoms appeared to deteriorate during that initial stage, a repeat challenge stage was offered and an additional repeat stage was offered if those results were inconclusive. The assessment of behavioral response was rather basic and limited to two types of subjective unstructured assessments by parents: a daily diary of symptom scores and an overall assessment at the end of each challenge phase. The results were assessed for individuals and not for the group as a whole. However, no statistical analyses of the data were apparently made. Four of the 19 subjects had a pre-study history of adverse behavioral reaction to yellow color but none of these four showed any adverse effects with the controlled color challenge. Among the remaining 15 children only 3 showed any adverse effects of challenge and only one of these 3 "responders" developed adverse behavioral symptoms that were associated with the preservative challenge but not with the color challenge. Confidence in the reliability of these study results, however, is lessened by the questionable accuracy and sensitivity of the limited and unstructured behavioral assessment using a very limited, poorly defined and non-validated scoring system, and the conduct of this assessment only by the parents with no other source of assessment to confirm the parental observations. Additional limitations include no apparent verification of the effectiveness of the blinding for the parents and children (blinding of drinks with tartrazine is reportedly very difficult due to the intense coloration of the dye), and no description of statistical analyses of the data. Even considering these caveats, this study does suggest that anecdotal reports of definite adverse behavioral (and other) intolerance to artificial colors, specifically yellow colors, are not clearly evident when tested under controlled clinical conditions.

- **Pollock and Warner (1990)** conducted a placebo controlled artificial color challenge trial of children specifically identified by parents as exhibiting food additive intolerance, whose behaviors reportedly improved on a diet excluding food/color additives and returned (primarily poor concentration and excess fidgetiness) shortly after consuming even small amounts of food additives. The 19 children used for the study, who were heterogeneous for hyperactivity but with problem behaviors, were maintained on their additive-free diets throughout the 7 week trial period. On 2 separate weeks children were administered a daily gelatin capsule containing a 125 mg blend of artificial food colors; placebo capsules with lactose were given during all other weeks. Parents were blind to treatment order but were not asked whether they could differentiate the color and placebo capsules. Two parent-based outcome measures were used: (1) a daily Conners behavior rating supplemented for somatic/allergy symptom assessment and (2) a weekly overall behavior assessment. The parent daily behavior ratings showed significantly enhanced problem behaviors during the color challenge weeks compared with placebo weeks and no treatment order effects. However, it should be noted that these were very small differences between the color and placebo weeks. The daily somatic/allergy scores showed no treatment related differences, indicating that the changes in behavior scores were not related to or secondary to somatic symptoms. Additional analyses showed no difference in behavior ratings between week days 1 and 7, indicating that the treatment related effects of food color intolerance began occurring after the first challenge and did not result in cumulative effects. However, the absence of any treatment related effects on behavior based on the parents' overall assessments at the end of each week provides no supportive confirmation for the color effects seen with the parent daily behavioral ratings. Unfortunately, the use of only parents' observation of behavior to measure treatment related effects does not enable independent sources of information to help resolve this inconsistency. While this inconsistency may reflect a questionable reliability of one of the measures, it seems more likely that the very small treatment effect was simply not detectable in the weekly assessment. However, since the claim by parents of being able to detect behavioral deterioration after consumption of even small amounts of food additives was a criterion for entry into the study, the fact that the parents weekly assessment could not detect an overall behavioral change even after daily challenges with a high dosage level of 125 mg artificial colors and that rather small behavioral differences were revealed in the parent daily ratings suggest that there is little substantial or functionally (biologically) relevant association between artificial colors and problem behaviors in children reported to be food additive intolerant.

- **Sarantinos et al (1990)** conducted a placebo controlled challenge trial of two artificial food colors, tartrazine and sunset yellow, on the behavior of children diagnosed with attention deficit disorder (ADD). *(It should be noted that only an abstracted summary of this study was available for review).* Of the 13 ADD children recruited for this study, 9 were considered responsive to a color-free diet and 4 were considered uncertain diet responders. Several children were reported to be intolerant to multiple food items and, oddly, two were claimed to be behaviorally intolerant to orange juice, which was the vehicle used in this study. The children were maintained on a color-free diet during the study and assigned to one of two treatment groups. Over the 28-day study period, the children in Group 1 (n=6-7) were challenged with an orange drink containing only 10 mg tartrazine on 6 random days. The children in Group 2 (n=6 or 7) were also given color challenges on 6 random days but on 3 days the orange drink contained 10 mg tartrazine and on the 3 other days the orange drink

contained 10 mg sunset yellow. On all non-color challenge days, children received a placebo drink. The only source of outcome measure was the parent behavior ratings using two rating scales (Conners scale and Rowe Behavioural Inventory scale). No procedure to verify effectiveness of blinding for the parents was described. One unidentified subject did not complete the trial. Based on repeated measures analysis of variance of the parent behavioral ratings, 2 of the 6 or 7 ADD children in Group 2 showed significant adverse changes in behavior (i.e., increased irritability, impulsivity, restlessness, and sleep disturbance) associated with both tartrazine and sunset yellow challenges. Both children were atopic and, among those whom parents had reported to be responsive to color-free diets, one child was intolerant to multiple food items whose mother could correctly identify color or placebo days (25/28 times but unclear whether this was based on child's behavior or broken blind). It was not stated whether either of these two responders were the 2 children reported to be intolerant to orange juice. Although not specifically stated, it appears that the 6 or 7 children in Group 1 did not show any significant response to their tartrazine challenges, thereby providing no confirmation or support of the responses noted in the two Group 2 children. Overall, based the absence of other independent outcome measures to corroborate the parental ratings, the inconsistency in findings between Group 1 and Group 2 children, the uncertainties regarding the effectiveness of blinding and the limited details about the experimental design, there are considerable questions about the confidence in reliability and biological significance of the findings, precluding any credible interpretation of these findings with regard to the possible association of the artificial colors, tartrazine and sunset yellow, and adverse behavioral effects in children with ADD.

- **Rowe and Rowe (1994)** uniquely conducted a double-blind, placebo controlled, repeated measures clinical challenge trial providing a full dose response assessment of tartrazine (Yellow 5) to establish whether there is an association between this artificial color and behavioral change in children with suspected hyperactivity. The first stage of this study was an open trial to identify children suspected of hyperactivity who might be sensitive to artificial colors based on parents reporting behavioral improvement on a color-free diet and deterioration with ingestion of foods containing artificial colors. Behaviors consistently reported related to irritability, restlessness, and sleep disturbance (behaviors not typically characteristic of hyperactivity syndrome) (also reported in Rowe, 1988). A total of 34 children with suspected hyperactivity, categorized as *likely* diet responders (23) or *uncertain* diet responders (11), and 20 children without problem behaviors participated in the challenge stage of this trial. All subjects were maintained on a color-free diet and given one double capsule (children ≥6 years) or packaged orange juice (children <6 years) on each morning of the 21-day study period. Over 6 random days, each child was administered (in capsule or orange juice) all six dosages of tartrazine (1, 2, 5, 10, 20, and 50 mg) once with one dosage per day in random order, except that the 20 and 50 mg dosages were administered toward the end of the study period. A placebo washout period of at least 2 days separated each challenge day for each child. Other than use of specially designed double capsules with an inner capsule surrounded by lactose, no description was given as to blinding procedures or to verification of effectiveness of blinding. Daily parent ratings provided the only outcome measures of behavioral change. The primary rating scale was a behavioral inventory devised by the investigators and validated to discriminate between color and placebo ingestion. A parent Conners rating was also included for comparative purposes only but not analyzed for treatment effects. The parent ratings identified a total of 24 of the 54 children tested as showing consistent behavioral reaction to the tartrazine challenges, significant at all six

dosage levels in clearly dose related fashion with amplitude and duration of effect increasing with increasing dosage. Beyond 10 mg there appeared to be a ceiling effect in magnitude but higher doses increased duration of effect. Most tartrazine reactors were among the children categorized as "likely diet responder" (19/23), while only few tartrazine reactors were in the "uncertain diet responder" group (3/11) and even fewer among the control subject with unknown diet sensitivity (2/20). There were no significant sex or age related factors and, while all tartrazine reactors were atopic but none showed an allergic reaction to tartrazine, it was not stated whether the nonreactors showed similar results. The principle behavioral features reported to be associated with tartrazine exposure, based on the devised behavior inventory, were irritability, restlessness, and sleep disturbance. It is possible that the devised behavior inventory rating may underestimate changes in attention behaviors. The two basic experimental issues that potentially affect the confidence in the reliability of the behavioral findings in this study include the adequacy of the study blind and the use of parent ratings as the sole measure of behavioral change. With regard to blinding, while the study was conducted under double-blind conditions, there was no apparent effort to determine whether the difference between placebo and color capsules and orange juice were adequately blinded or to verify the effectiveness of the blinding procedures for the parents and children. However, with the results showing clear dose response effects, it seems likely that the integrity of the blind was adequately maintained. With regard to the use of parental ratings as the sole outcome measure in this study, confidence in the reliability of study findings would certainly be strengthened if other measures of behavior were also used to corroborate the parental findings. However, in view of the fact that the parent ratings were able to discern significant incremental differences in behavior and in clear dose related fashion, there appears to be sufficient confidence in the reliability of these study findings based solely on the parental behavioral ratings. Overall, these findings do indicate that intolerance to tartrazine can cause adverse behavioral effects, not necessarily associated with hyperactivity, in selected subgroups of diet sensitive children with problem behaviors and possibly in a small subgroup of control children without problem behaviors and unknown diet sensitivity.

- **Bateman et al (2004)** designed a double blind, placebo controlled challenge trial to determine whether food additives can cause hyperactive behaviors in preschool children sampled from the general population with and without hyperactivity and with and without atopy (allergic sensitivities). The 277 preschool children, who completed the challenge trial, were not diagnosed ADHD but were assessed for hyperactivity based on parent behavior ratings and were designated hyperactive (HA) or non-hyperactive (non-HA). All children were further assessed for atopy (allergy sensitivity) and identified as either atopic (AT) or non-atopic (non-AT). They were then sorted into four test groups: HA/AT (36), HA/non-AT (75), non-HA/AT (79), and non-HA/non-AT (87), each group having slightly more males than females. The children were not reported diet responsive but were placed on a diet free of artificial colors and preservatives throughout the 4 week study period. During two weeks of that period each child received a daily drink of either the placebo fruit juice or the challenge fruit juice containing a 20 mg blend of artificial colors plus 45 mg sodium benzoate (one type of treatment per week determined randomly). Two weekly behavioral outcome measures were used: (1) a clinically based aggregate test hyperactivity score (ATH), validated as distinguishing between hyperactive and non-hyperactive preschool children at baseline, and (2) an aggregate parent hyperactivity rating score (APHR). Analyses of variance and co-variance were used to analyze the data. Based on grouped data analyses, APHR (parent ratings) scores showed an overall significantly greater increase in hyperactivity behaviors for

the active color/benzoate challenge than for the placebo with no difference in response between the groups of children categorized HA and non-HA or between the AT and non-AT groups. Curiously, both placebo treatment and color/benzoate challenge increased levels of hyperactivity behaviors, but a statistically greater increase in hyperactivity behaviors occurred overall during the active color/benzoate challenge week. Group analysis of the ATH (clinical measures) scores, however, showed no significant behavioral effects associated with the color/benzoate challenges. The fact that the ATH score provided no confirmation or support for the APHR findings is particularly relevant in assessing the significance of the ATH findings, since validation of the APHR scoring was based on distinguishing between hyperactive and non-hyperactive preschool children at baseline. An additional issue to be considered in evaluating these contradictory findings is the effectiveness of the blind for the parents. While efforts were made to maintain the blinding conditions for this study, it was reported that approximately half of the parents could correctly identify the treatment but it was not clear whether this was due to behavioral changes in the children or to detectable differences between placebo and challenge drinks. In view of the contradictory findings between the two primary outcome measures, together with the uncertain effectiveness of the blind, there is limited confidence in the reliability of this study's findings. This study provides no evidence that food additives cause hyperactivity in preschool children from the general population. At best this study may suggest that non-allergy related intolerance to artificial food color additives and benzoic acid may occur in some children with or without hyperactivity related behavior and that this intolerance may be associated with only limited behavioral changes.

- **McCann et al (2007)** conducted a controlled challenge trial to test whether artificial food colors and additives (AFCA) affected childhood behaviors in a general population of preschool and school age children and in part to replicate a previous study (Bateman et al, 2004). The study sample of 153 preschool (3-4 years) and 144 school age (8-9 years) children were recruited from the general population, with neither group being diagnosed hyperactive or having special problem behaviors. The challenge material was a fruit juice drink containing one of two different color blends plus sodium benzoate (Mix A and B). Two sets of each Mix were used, one containing lower levels of colors (20 mg Mix A and 30 mg Mix B) to challenge the preschool children and the second set containing higher levels of colors (25 mg Mix A and 62 mg Mix B) to challenge the school age children. All challenge mixes contained 45 mg sodium benzoate. After baseline, children were maintained on a diet excluding artificial colors and sodium benzoate throughout the remaining 6 weeks of the study. During each of 3 different weeks, children were given daily treatments with either Mix A, Mix B or placebo fruit juice, with one type of treatment per week; the order of treatment across weeks was determined randomly for each child. While all study participants were blind to treatment and other efforts were made to ensure the blinding of the study, effectiveness of blind for parents and children was not verified. The sources of outcome behavioral measures included: (1) teacher ratings, (2) parent ratings; (3) trained observer classroom observations; and (4) a computerized continuous performance task was used as an additional objective source measure only for school age children. However, for purposes of determining whether behaviors were affected by either of the color/benzoate treatments, the data from all sources of behavioral measures were combined for each age group separately into a single composite Global Hyperactivity Assessment score (GHA) which was then analyzed to identify any significant treatment related behavioral effects. Based on differences in the composite GHA scores for the full sample of children, the young preschool children

responded to Mix A challenge with significantly elevated hyperactivity scores but did not respond to Mix B challenge. Contrarily, the older school aged children responded to Mix B challenge with significantly elevated hyperactivity scores but did not respond to Mix A challenge. These differences in response to the two mixtures of colors/benzoic acid seem to be a puzzling inconsistency and the investigators even acknowledge the need for additional study" to establish whether the age-related differences seen in the present study can be replicated". While the specific nature of the behavioral effects are not clear (none of the component outcome behavior measures were described), whatever behavioral changes may have occurred were apparently of rather low magnitude (effect size of 0.18). This would suggest that the type of treatment effects reported in this study, even though the investigators referred to increases in levels of "hyperactivity", were not the disruptive excessive hyperactivity behaviors of ADHD but more likely the type of overactivity exhibited occasionally by the general population of preschool and school age children. However, due to the study weaknesses and caveats of this study, particularly the puzzling inconsistency in response by the two age groups of children and the rather small treatment related changes, the reported effects appear to be equivocal and of uncertain biological relevance. This study does not provide credible evidence of deleterious effects of artificial food colors and additives on children's behavior nor does it lend any support for the contention that food additives exacerbate hyperactive behaviors (inattention, impulsivity, overactivity) in children.

Group I: Summary Conclusions

Pre-1982 Clinical Trials (Summary Conclusions)

In attempting to summarize the conclusions that may be drawn from the Group I studies, particular note should be given to an NIH Consensus Development Conference on Defined Diets and Childhood Hyperactivity convened in 1982 to assess the status of extant information relating defined diets (specifically, the Feingold diet) and hyperactivity disorders in children (NIH, 1982). At that time, at least 16 controlled clinical trials (reviewed in this project report and color-coded grey† in Table 1), two diet crossover trials and 14 specific challenge trials, had been conducted to assess the Feingold hypothesis. Based on the committee's review of those trials and other information, the NIH consensus report alluded to differences and inadequacies in the designs of the controlled clinical trials making analysis difficult, but concluded that the available studies "*did indicate a limited positive association between 'the defined diets' and a decrease in hyperactivity...that involved only a small proportion of patients; furthermore, the decreases in hyperactivity were not observed consistently. Studies indicated that some hyperactive children on a defined diet experienced an increase in hyperactivity when given moderate doses of artificial food colors.....by only a small group of patients and the increase was not consistently reported by teachers, parents, and other observers*". This highly qualified conclusion fairly represents this reviewer's assessment of the information from the Group I clinical trials conducted prior to 1982 and refers to several of the same types of problematic and confounding issues identified in this reviewer's evaluation of most of those same early clinical studies. These issues include varying degrees of inconclusive or equivocal findings (Swanson and Kinsbourne, 1980), only a small percentage of responders under controlled conditions (Conners et al, 1976; Harley et al, 1978; Weiss et al, 1980), inconsistent or unconfirmed reports of treatment effects between different sources of behavior assessment (Goyette et al, 1978; Rose, 1978; Spring et al, 1981; Swanson

and Kinsbourne, 1980; Weiss et al, 1980;Williams et al, 1978; Mates and Gittelman, 1981) or due to inexplicable treatment order effects (Conners et al, 1976; Harley et al, 1978), and questionable effectiveness of blinding conditions (Conners et al, 1976; Harley et al, 1978; Goyette et al, 1978; Levy et al, 1978; Levy and Hobbes, 1978; Conners, 1980; Swanson and Kinsbourne, 1980; Weiss et al, 1980). In view of these caveats, the treatment related findings from these pre-1982 studies were generally considered by this reviewer as only suggestive of limited beneficial effects of Feingold's diet in hyperactive children and a limited association between artificial colors, even at high daily dosages of 78 to 150 mg mixed colors (Swanson and Kinsbourne, 1980; Mattes and Gittelman, 1981), and adverse behaviors in a small subgroup of children with hyperactivity or other problem behaviors. Subsequently, in 1983 a meta-analysis of trials, including those reviewed at the 1982 NIH Consensus Conference, focused on the Feingold hypothesis, particularly as it related to the effectiveness of his defined diet in treating childhood hyperactivity (Kavale and Forness, 1983). These authors concluded that their analysis of trials up to that time point provided no support for the Feingold hypothesis and that his dietary treatment was of questionable effectiveness, producing only slight improvements in behavior of hyperactive children with little substantive changes to the basic elements of attention and learning.

Post-1982 Trials (Summary Conclusions)

Following the 1982 NIH Consensus Conference and up to the present time, at least 10 additional clinical trials dealing with the association between food colors and problems behaviors related to hyperactivity of childhood were conducted (Table 1). In summarizing the conclusions that may be drawn from these 10 post-1982 clinical trials, it should be noted that, comparable to the pre-1982 trials, many of these trials were found to have varying limitations to their experimental designs which resulted in varying levels of confidence in the reliability of their study findings. Among the study limitations were included: use of unstructured non-validated rating systems (David, 1987; Wilson and Scott, 1989), non-blinded study personnel responsible for behavior ratings (David, 1987) or effectiveness of blinding not verified or questionable (David, 1987; Thorley, 1984; Rowe, 1988; Wilson and Scott, 1989; Pollock and Warner, 1990; Sarantinos et al, 1990; Rowe and Rowe, 1994; Bateman et al, 2004; McCann et al, 2007), use of single source for behavioral outcome measures not confirmed by other sources of measurement (Rowe, 1988; Wilson and Scott, 1989; Pollock and Warner, 1990; Sarantinos et al, 1990; Rowe and Rowe, 1994), no randomized assignment to treatment (David, 1987), and no statistical analysis (David, 1987; Wilson and Scott, 1989) or incomplete presentation of data (David, 1987; Wilson and Scott, 1989; Sarantinos et al, 1990; McCann et al, 2007). Across trials, the reported findings of adverse reactions to color additive challenge were mixed (4 negative, 3 inconsistent, 3 positive), as were the levels of confidence in those various trial results. Among the four challenge trials showing no adverse responses to color challenge, there was reasonable confidence in the results from one trial showing no adverse behavioral or cognitive effects of 92 mg mixed colors in retarded children (Thorley, 1984), moderate confidence in another showing no effect of Red 40 on heart rate or EEG activity a hyperactive children (Salamy et al, 1982), and low confidence in two trials with tartrazine, sunset yellow and preservatives (David, 1987; Wilson and Scott, 1989). For the three trials showing inconsistent or inconclusive color effects, there was reasonable confidence in the results of one trial showing inconsistent but very small behavioral effects of daily 125 mg mixed color challenge in problem behavior children (Pollock and Warner, 1990), moderate confidence in one trial with inconsistent behavioral effects with 2 mixes of 20-30 mg artificial colors in preschool and school age children from the general population (McCann et al, 2007), and limited confidence in one trial with inconsistent small

behavioral effects and uncertain effectiveness of parent blinding with 20 mg mixed color challenge in pre-school children from the general population (Bateman et al, 2004). And, among the three trials associating adverse behavioral effects with artificial color challenge, there was reasonable confidence in one study showing clear dose response behavioral effects at each challenge dose of tartrazine from 1 to 50 mg primarily in problem behavior children but also in several children without behavioral problems (Rowe and Rowe, 1994), moderate confidence in one trial reporting 2 children (one hyperactive, the other non-hyperactive with problem behaviors) responding to 50 mg tartrazine or carmoisine with adverse behaviors that may or may not be characteristic of hyperactivity (Rowe, 1988), and low confidence in the third trial reporting several children responding to intermittent challenge with 10 mg tartrazine and sunset yellow (Sarantinos et al, 1990).

Generally, the various reported findings across these 10 reviewed post-1982 portion of Group I trials, suggests that certain susceptible subgroups of problem behavior children with and without ADHD and, possibly, certain susceptible children from the general population without particular behavioral problems may exhibit a unique intolerance to artificial food colors resulting in typically small to moderate adverse behavioral changes which may not necessarily be characteristic of the ADHD syndromes. A relatively recent meta-analysis of clinical trials dealing with artificial food colors and hyperactivity in children (Schab and Trinh, 2004) was conducted and considered the relevant artificial food color challenge studies conducted through 2004, including most of the Group I trials reviewed in this project report. Their analyses across all trials found small but significant treatment effects for the color challenges, suggesting an association between ingestion of artificial food colors and hyperactivity behaviors. Their secondary analyses also implicated artificial food colors more in provoking general behavioral disturbances than hyperactive symptomatology and suggested that sensitivity to artificial food colors may not be limited to those children with clear-cut hyperactive syndromes.

For many of the color challenge trials the children recruited for that study were specifically selected in part because of their reported diet responsiveness, showing improvement in their hyperactive and problem behaviors on an open (non-blinded) defined diet (usually Feingold's diet) which at least excluded artificial food colors, and their dramatic deterioration of behavior after exposure to foods containing artificial colors. However, in the controlled challenge phase of testing, typically few, certainly not all, of the subjects were found to exhibit adverse behavioral reactions to the controlled color challenges. Speculatively, for those subjects who did not respond to the controlled color challenges, part of their reported diet responsiveness and dramatic sensitivity to foods with artificial colors (to the extent that non-blinded reported responses were real) may have been associated with a sensitivity or intolerance to certain food items/components other than artificial colors (Stare et al, 1980). This possibility of sensitivity or intolerance to a broader spectrum of food substances, other than artificial colors, is the focus of the Group II trials also reviewed in this project.

Group I Trials: General Conclusion

The focus of the 26 Group I trials was to assess the validity of Feingold's hypothesis relating food additives and ADHD and other problem behaviors in susceptible children. The clinical trials were typically designed as double blind, placebo controlled studies either to assess the effectiveness of the Feingold diet as a dietary treatment of ADHD or to test the role specifically

of artificial colors in causing or provoking adverse behaviors related to ADHD or other problem behaviors in children. Various caveats in the study designs and uncertainties in the study results make interpretation of the study findings difficult and limited in scope. Within this context, the results from the Group I controlled clinical trials, overall, suggest the conclusion that certain subgroups of children with problem behaviors that may or may not be related to ADHD/ADD and, possibly, certain children from the general population without particular behavioral problems may exhibit a unique intolerance to artificial food colors resulting in typically small to moderate adverse behavioral changes which may not necessarily be characteristic of the ADHD syndromes.

Group II Trials (n=7)

As noted previously, these controlled trials focused on a broader array of substances by assessing the role of food intolerance, i.e. adverse effects of food itself, in hyperactive and problem behavior children. Of the 7 trials in this group, 2 were "diet crossover trials" both of which used a particular type of elimination diet (i.e., oligoantigenic diet, few foods diet, and Alberta Children's Hospital diet) that excludes all foods, additives, including artificial colors, and food components assumed to provoke adverse behavioral reactions in certain children and assessed the effectiveness of these diets in improving the behavior of hyperactive or problem behavior children. The remaining 5 trials were "*specific challenge trials*" which assessed the adverse behavioral and other effects of various suspected provoking food items (pre-identified for individual subjects under non-blind conditions) in children diagnosed ADHD or with problem behaviors. The consideration in these Group II trials of a broad range of food substances that may be associated with hyperactivity appears to be responsive to the concern expressed in the 1982 NIH Consensus Report that "controlled challenge studies have primarily involved the administration of food dyes in children" and "do not appear to have addressed adequately the role of diet in hyperactivity".

Typically, the procedure for Group II "*diet crossover trials*" was basically the same as for the Group I trials, i.e. randomly assigning a group of children to either a defined elimination diet or a matched placebo diet for a specified period of time, then crossing them over to the other diet condition for an equivalent period of time and evaluating the behavioral responses to both diet conditions. The "specific challenge trials" were typically conducted in three phases: Phase I – identify a study group of children who reportedly show improved behavior when placed on an open (non-blind) food elimination diet; Phase II – methodically re-introduce foods/components to the 'food elimination diet' in non-blinded fashion to tentatively identify specific provoking food items for individual subjects (foods not producing a reaction were kept in the diet on which the child was maintained throughout the subsequent challenge phase of testing); Phase III – conduct double blind, placebo controlled challenge testing with one or more suspect provoking food items to verify and assess their adverse behavioral effects under controlled conditions.

Individual Trial Summary Reviews (Group II)

- Based on the possibility that some type of food intolerance (food allergy) may play a role in causing or aggravating the hyperkinetic syndrome in children, **Egger et al (1985)** conducted one of the first trials to treat problem behavior children with an oligoantigenic diet. In a clinic setting 76 children, considered atypical hyperactive many with neurological disorders,

allergic conditions and other physical conditions, were placed on an open (non-blinded) oligoantigenic diet and 62 (82%) were reported to show improved behavior. Skin-prick testing showed that atopy status (allergic sensitivity) did not differ between the children who did and did not respond to the diet, but IgE levels were higher in the responders. Twenty-eight of the "food intolerant" children (average age approximately 7 years) were tested in open re-introduction trials to identify specific foods/components to which they were sensitive. While all children appeared to react to more than one food/component, the most common suspect provoking food ingredients were benzoic acid and tartrazine. The varieties of food items subsequently used as challenges included cow milk, orange juice, wheat, tinned food, tartrazine (150mg) in capsule, or benzoic acid (150mg) in capsule. Eight of the 28 subjects received either tartrazine or benzoic acid as their challenge. Placebo items were matched to blind the challenge. All participants were blind to order of challenge and staff members found no difference in taste between active and placebo treatments, but families were not assessed to verify the effectiveness of blinding. In the double blind, placebo controlled phase of this trial, each child was randomly assigned to receive either his/her suspect provoking food item or placebo for 1 to 2 weeks and an additional 1 to 2 weeks with the opposite treatment, with at least a 2-week washout period between treatments. Multiple outcome measures were used, including parent Conners behavioral ratings, overall behavior assessments by parents, clinician, and psychologist, and a battery of psychological tests. Under these controlled challenge conditions, the overall behavior assessments by parents, psychologist and clinician linked better behavior with the placebo challenge significantly more often than with the suspect provoking food challenge. Parent Conners behavior ratings also showed significant treatment effects, with the suspect provoking food challenge eliciting more hyperactive behavior than placebo. While group analyses linked the provoking food items with adverse behavioral changes, graphic representation of some of the challenge data showed that approximately 6 children did not respond adversely to their challenge. It is not known with what food items these children were challenged. In contrast to the subjective rating and assessment measures, none of the objective psychological tests showed any significant adverse behavioral effects with the suspect provoking food challenge, although non-significant trends did generally indicate placebo behavior better than active challenge. Since significant adverse effects of provoking food items was corroborated by several sources of outcome measures (parents, clinician and psychologist), it seems likely that the failure to detect treatment related effects with the objective psychological test in this study may have been due to the possibility that those tests were not sensitive enough to detect either the type or magnitude of behavior changes specifically associated with food intolerance. Although there is some uncertainty regarding validation of the effectiveness of the blind for the parents and children, this study does provide sufficiently reliable findings to indicate that some children with problem behaviors may be intolerant to a variety of food items, including but certainly not limited to artificial food colors or preservatives.

- **Kaplan et al (1989)** designed a crossover dietary trial to assess the effects of a special foods elimination diet (Alberta Children's Hospital Diet, ACH) specifically in preschool hyperactive male children in day-care with sleep problems and/or allergy-type somatic symptoms. Twenty-four preschool ADHD children (3.5 – 6 years), reported by parents to have sleep or physical problems participated in this study. No stimulant medication was used during the study. The experimental diet, ACH, eliminated artificial food colors, flavors, preservatives, MSG, chocolate, caffeine, and milk, dairy products, and natural salicylates for some children, and decreased simple sugars. An apparent equivalent diet was used as

placebo. Extensive efforts were carried out to ensure effectiveness of the blind. All children were given the placebo equivalent diet for 3 weeks and the experimental ACH diet for 4 weeks, with these two diet treatments given in randomized counter-balanced order across subjects. While multiple sources of behavioral outcome measures were planned, including parent ratings/records (included Conners rating scale supplemented with individualized problem behaviors identified by parents), day-care worker ratings (Conners rating scale not individualized), independent observer records, and laboratory psychometric testing, only parent ratings/records were completed. Day-care worker ratings were only partially completed (complete data for 10/24 subjects) and the other two planned measures were terminated due to technical difficulties or discontinued due to subjects being deemed "untestable", respectively. Other non-behavioral testing (blood chemistry/hematology/basic physical) was completed. Statistical analysis of data was conducted using repeated measures multivariate and univariate ANOVA with Tukey's method of multiple comparisons. Based on the parent Conners behavior ratings without the individualized item, there was a small but statistically significant improvement of the children on the ACH diet (mean score 10.8) compared with the equivalent placebo diet (13.1). Evaluation of the parent Conners scores for individual children revealed that not all children were equally responsive to the ACH diet. Compared with the equivalent placebo diet, 14 (68%) of the 24 children showed some behavioral improvement on the ACH diet, with 9 showing at least a 25% improvement and 5 showing a milder improvement. The remaining 10 children (42%) were unresponsive to dietary intervention. The parent ratings of the individualized problem behaviors and the sleep records showed comparable overall improvement on the ACH diet (the numbers of individual responders were not identified). The limited the day-care worker ratings data showed marginal but non-significant diet related differences in behavior. Also, the physical signs/symptoms, clinical or nutritional measures showed no reliable diet related differences. Overall, since the day-care worker ratings, although completed on less than half of the subjects, did not confirm a significant diet related effect and since no other source of outcome behavioral data was available to resolve this inconsistency, the diet related findings in this study are considered inconclusive but suggestive that a special food elimination ACH diet may result in some level of behavioral and sleep improvement in some hyperactive pre-school children. While this suggests that some hyperactive children may be intolerant to some dietary elements, it is unknown what specific food item(s) or component(s) might be involved.

- Based on the premise that food intolerance may be associated with adverse behaviors in certain ADHD children through an allergic/immunologic mechanism, **Egger et al (1992)** designed a placebo controlled clinical challenge trial to determine whether hyperkinetic children can be desensitized to the adverse effects of provoking foods using an enzyme-potentiated desensitization treatment (EPD). In initial non-blinded phases of this study 116 out of 185 hyperkinetic children (63%) were reported to show improved behavior on an oligoantigenic diet and, subsequently, specific provoking foods/components were identified for each child based on recurrence of hyperkinetic behaviors or other associated physical symptoms. Chocolate, artificial colors, cow milk, egg, citrus, wheat, beet sugar and nuts were among the most common provoking food items. Forty of these food intolerant children participated in the double-blind, placebo-controlled trials and were maintained on an oligoantigenic diet throughout the study period. Half of the children received injections of either placebo/buffer solution or an EPD solution of mixed food antigens (multiple foods, additives, colors, and preservatives). Both solutions were reported to be "colourless" and

parents and children were blind to treatment. The effectiveness of the experimental treatment was determined by reintroducing provoking foods individually and assessing the children's response. Parents maintained a record of hyperactivity behaviors and other symptoms (e.g., headache, abdominal discomfort, bloating, and diarrhea). Treatment effectiveness was based on two measures: number of provoking foods given before parents stopped food reintroduction when adverse symptoms occurred and whether parents rated treatment as successful in preventing or reducing symptoms from one or more provoking foods. Statistical analyses used the Fisher's exact test. Atopy and IgE levels were determined. Based on the measure of parents stopping the reintroduction of provoking foods, adverse response to provoking food items occurred less frequently in the EPD treated group than in the placebo group. Specifically, 15 of the 20 children given EPD treatment were able to eat previously identified provoking food items without adverse reactions, but only 7 out of 20 placebo treated children were able to do so. Also, more parents of children treated with EPD than those treated with placebo thought the treatment was successful. While both behavioral and physical symptoms were reportedly used to assess response to provoking food items, the physical symptoms, such as abdominal discomfort, bloating and diarrhea, were usually the first to appear but no description of behavioral effects was included. These results were not related to atopic status or IgE levels in these children. Several issues involving experimental design weaknesses affect confidence in this study's finding. One issue is the questionable masking of the difference between the placebo and EDP solutions which is central to the study blind and the reliability of the subjective parental observations. The presence of 6 different food colors in the EDP solution seems to be inconsistent with the EDP solution being described as "colourless". Any notable difference in coloration between the placebo and EPD solutions would make effective blinding of the parents and children to treatment questionable and introduces a possible unintentional assessment bias. A second study limitation is the fact that no other source of information was available to confirm or support the parent detection of adverse response to the provoking foods items. A third issue involves the fact that no structured or validated rating system was used to assess the behavioral changes and it is unclear what types of adverse behavioral effects occurred. In view of these caveats, particularly the questionable masking of treatment and the use of parent reports as the sole outcome source, there is limited confidence in the reliability of the study data. Overall, however, the reported findings may be viewed as suggestive that intolerance to a variety of food items, associated with physical, and possibly behavioral, symptoms in some hyperkinetic children, may be lessened by a process of desensitization which further suggests some immunologic process.

- In an effort to determine whether anecdotal reports of food intolerance associated with ADHD can be verified under controlled clinical conditions, **Carter et al (1993)** conducted a double blind, placebo controlled clinical trial testing the effects of a "few foods" elimination diet in hyperactive children. The "few foods" diet is analogous to the oligoantigenic diet in that both eliminate every food or additive assumed to provoke behavioral reactions in certain children. In the open (non-blind) phase of testing 59 out of 78 of the ADHD children tested (76%), many of whom were already on some restricted diet, were found to be food intolerant based on their reported behavioral improvement on the "few foods" elimination diet, which was continued throughout the duration of this study. In the open reintroduction phase a large number of food items were identified that provoked an adverse behavioral response sometimes with physical symptoms. Nineteen of the diet responsive ADHD children completed the main controlled challenge phase of testing. Most children were challenged

with a series of more than one provoking food item, each disguised in a tolerated food which served as the placebo. Artificial colors (daily maximum of 26 mg) were given in capsules (glucose as placebo). The number of subjects challenged with various food items was not specified. Sufficient measures were taken to ensure blinding of study. Each child was challenged with a suspect provoking food item or the placebo each daily for one week in randomized order with a two-week washout period between active and placebo treatments. The multiple outcome measures included parent and psychologist ratings and 2 laboratory tests (paired associate/learning and familiar figures/impulsiveness). The parent ratings showed statistically significant deterioration of behavior with the provoking food item challenges compared with the placebo, but the magnitude of the behavioral change was small and the behaviors most affected related to irritability rather than attention deficit. The psychologist ratings also showed challenge with the provoking food items to significantly increase behavioral response, primarily fidgetiness. The laboratory test of impulsiveness also showed significant adverse food challenge effects but the learning task was unaffected. While all of the treatment related behavioral changes were in the direction of negative effects, the small magnitude of these changes makes their biological relevance unclear. The possibility of synergistic or additive effects from multiple provoking food items represents a data gap and should be addressed in subsequent studies. Notably also, the types of behaviors elicited related more to irritability and fidgetiness than on attention deficit, suggesting behavioral changes not necessarily associated only with the hyperactivity syndrome. Overall, this study presents credible findings indicating that parent anecdotal reports of a selected group of ADHD children showing improved behavior in a non-blinded trial of a "few foods" elimination diet can be confirmed to some extent in controlled clinical trial. These findings also support the conclusion that some ADHD children may have intolerances to various foods/components, including artificial food colors and additives, which may cause or exaggerate adverse problem behaviors which may or may not be related to the hyperactivity syndrome. However, with the use of general food items in this study it is not possible to identify specific food chemicals that may be causing the adverse effects.

- **Boris and Mandel (1994)** focused on the role dietary components may play in ADHD of childhood by conducting a double blind, placebo controlled food challenge trial to determine whether reported responses of ADHD children to suspect provoking food items were reproducible under controlled test conditions. Prior to the controlled challenge phase of the study, non-blinded testing was conducted to identify ADHD children who showed improved behavior on an open food elimination diet, comparable to the 'few foods'diet, (19 out of 26 children responded to diet with responders having parent mean Conners score of 25 prior to diet and 9.6 on diet) and who responded adversely to certain food items. More atopic children than non-atopics responded to the elimination diet, contradicting an earlier study which found no involvement of atopy (Egger et al, 1985). Sixteen of these "diet responders" (7.5 years) then completed the controlled food challenge trial. Each child was challenged with only the most provoking food item for that child, given as 5 grams powdered food or 100 mg blended artificial colors masked in a lentil soup or cranberry sauce. Provoking challenge items included milk (5), colors (4), corn (3), wheat (2), soy (1), and oranges (1). All study personnel were blind to treatment but effectiveness of blinding for parents and children was not verified. Note that no subjects reacted to preservatives in open challenge. During a 7-day experimental period, days 1, 2, and 7 were designated placebo days for all subjects. The food challenge and placebo treatment were randomly administered on days 3, 4, 5 and 6 with 13 subjects receiving food challenge two times, 2 subjects three times, and 1

subject one time. All other week days were placebo. Based on the parent behavior ratings (the sole outcome measure in this study), there was an overall significant increase in hyperactivity scores with the provoking food challenges (mean Conners score of 18.1) compared with placebo treatments (mean Conners score of 8.2). However, inspection of the graphed data show that only 11 of the 16 subjects actually responded to the provoking food challenge. The challenge food items for the challenge responders and non-responders were not identified. It should be pointed out that the noted difference in hyperactivity scores between the original (pre-study) diet period and the provoking food challenge period (mean Conners scores of 25 and 18.1, respectively) may be more apparent than real, since the blinding and treatment conditions under which the two ratings were conducted were completely different. Two experimental limitations of this study affect the confidence in the findings. Since the parent rating was the only measure of behavioral response, without any other sources of information to confirm or support the parent ratings, confidence in the reliability of these study findings is limited. A second limitation involves the study blinding which is critical to the integrity and reliability of any controlled study, particularly when subjective single source outcome measures are used. Although study participants were not informed of the challenge order and the placebo and food challenges could not be distinguished by taste, the effectiveness of the blind specifically for the parents and children was not validated leaving an uncertainty about the adequacy of the study blind. Overall, in view of these limitations, the confidence in the reliability of the study findings and their relevance is unclear. The findings may at least be considered suggestive that intolerance to various food components may cause or exaggerate hyperactive behaviors to a moderate degree in some ADHD children. The use of general food items does not allow specific food chemicals, which might be involved, to be even tentatively identified.

- **Schmidt et al (1997)** conducted a controlled diet crossover trial to examine the effectiveness of an oligoantigenic diet as a possible treatment for ADHD/Disruptive Behavior Disorder children and uniquely compared the diet with stimulant medication treatment. The study participants were an unselected sample of 49 children (6-12 years) hospitalized for treatment of ADHD/Conduct Disorder Behavior. The two dietary treatments were a restrictive oligoantigenic diet with beverages of apple juice and water, and a placebo diet of common food items with daily beverages containing 50 mg mixed artificial colors. All personnel were blind to treatment and diets designed so that differences in appearance would not allow identifying the experimental diet (this was not verified). All 49 children were tested with the experimental diet, but only 36 were additionally tested non-blinded with stimulant medication (methylphenidate, 0.4 mg/kg, po) for comparative purposes. After a 4-day baseline, children were given the oligoantigenic or placebo diets each for 9 days, with treatment order randomized. After a washout period of 9 days, children were treated with stimulant medication. Outcome measures consisted of both behavioral ratings (teacher rating in classroom and trained observer ratings at play and in laboratory during performance testing) and laboratory performance testing (continuous performance, paired associate learning, and activity). Notably, treatment effects were based only on the ratings data and did not include results from the objective laboratory testing. Blood levels of IgE were also measured but showed no difference between experimental diet and baseline conditions. Group level analysis showed significantly lower problem behaviors on the experimental diet compared with placebo but based only on the observer ratings at play and in the laboratory. The teacher ratings in the classroom showed no significant diet related differences in behavior. Analysis of rating data for individual subjects showed significantly lower problem

behaviors for 22/49 children (45%) in the laboratory setting and 21/49 children (43%) in the play setting. Twelve children (24%) were considered clinically significant "responders" (i.e.25% improved behavior on two measures). In the non-blind testing of methylphenidate treatment, 16/36 children (44%) were considered "responders". Out of 36 children who received both diet and drug treatment, 31% improved on drug only, 8% on diet only, and 14% on both. The findings in this study appear to be inconclusive due to the apparent discrepancy between observer and teacher ratings. Although several objective laboratory tests were conducted, none of those data were presented or analyzed and provide no information with which to resolve this discrepancy. In addition to the inconsistent findings between the two sources of behavioral assessment, the parents and trained observers, there are two other issues which may affect confidence in the results of this study. One issue involves the fact that no efforts were made to verify the effectiveness of the blinding for the children, i.e. whether or not they could actually tell which of the two diets was the restrictive oligoantigenic diet. The other issue is that, even though objective laboratory tests of behavior were actually conducted, none of those data were presented or analyzed. The investigators dismissed the data because they viewed these objective test data as less comprehensive, confounded by differences in motivation, insensitive to subtle treatment effects, providing redundant information, and not suitable for assessing response in a clinically relevant sense due to ceiling effects. The inclusion of objective laboratory tests to supplement the behavioral ratings was a feature viewed as adding credibility to the study by providing confirmatory evidence of treatment related effects. Dismissing these test results *post hoc*, without even presenting the data, lessens the ability to interpret the study findings that were presented. Overall, considering the limitations noted above, the results of this study are considered inconclusive and should be viewed only as suggestive of a possible limited beneficial effect of the oligoantigenic diet in certain children with a profile of disruptive behaviors including ADHD and Conduct Disorder or that food intolerance may exaggerate some adverse behaviors in a small group of select children with ADHD/Conduct Disorder.

- **Uhlig et al (1997)** designed a clinical challenge trial, using a special technique of EEG topographical mapping, to investigate whether changes in brain electrical activity occur in food sensitive ADHD children exposed to provoking foods. The 12 children (6-15 years) completing all phases of this study were outpatients from a special diet clinic and were diagnosed ADHD. In initial open (non-blind) trials, behavior ratings by parents and teachers showed improved behavior in these children on an oligoantigenic (few foods) diet and suspect food items provoking the return of adverse behaviors were identified for use in the challenge phase of this study. EEG mapping of each child was carried out under both conditions of "consuming" and "avoiding" provoking foods. For the "avoiding" condition, children simply maintained the oligoantigenic diet for 14 days before EEG mapping. For the "consuming" condition, children ate provoking foods daily for at least 5 days prior to and on the day of EEG mapping. During the EEG sessions two investigators also conducted Conners behavioral ratings. Notably, other than one investigator being blind to treatment order during evaluation of EEG recordings, no other study personnel particularly including parents and children were reported as being blinded to treatment during this study. Conners scores were significantly higher for all children under the "consuming" provoking foods condition compared with the "avoiding" condition (it should be noted that two other children who showed no challenge response to the provoking foods were excluded from the study). Significant EEG differences were also found between diet conditions, primarily in the fronto-temporal regions of the brain with few changes in the parieto-occipital areas. Relative to the

"avoiding" diet condition, when children consumed provoking food items, the percentages of fast beta-1 frequency electrical activity increased in frontal areas of the brain in conjunction with abnormal behavior ratings. The relative power spectra of all recordings while consuming provoking food items were significantly higher than during avoidance of the provoking food items. Unfortunately, in an otherwise well-conducted electrophysiological study, the absence of additional sources of behavioral measures to confirm the investigators' ratings lessens confidence in the reliability of the findings. Further, the critical absence of any apparent blinding conditions introduces the possibility of an unintentional influence on the behavior of the children and confounds interpretation of the EEG results and associated behavioral ratings. The children may have behaved differently not because of the diet but because they or their parents knew which diet they were receiving, which in turn could have affected the behavioral ratings and the EEG mapping. Also, the exclusion from the study of two subjects specifically because "there was no change in behaviour during the crossover study" is questionable. Overall, due primarily to the absence of blinding conditions and the use of a single source measure of behavioral changes, the results from this study cannot be interpreted as evidence of an association between exposure to "provoking" foods (food intolerance) and specific EEG changes or behavioral rating changes but may be suggestive of such an association. However, the results do appear to show credible evidence associating specific electrical activity changes in the frontal brain region with increased hyperactivity behaviors (indicated by increased Conners rating scores).

Group II Trials: General Conclusion

To summarize the conclusions from Group II trials, while all 7 Group II trials showed some findings indicative of food intolerance, 3 trials were inconclusive and there were varying levels of confidence in the findings for all trials. Among the 4 trials showing positive findings of food intolerance, there was a level of reasonable confidence in only one of those trials (Carter et al, 1993), and moderate or limited confidence in 2 trials (Boris et al, 1994; Egger et al, 1992) and low confidence in one trial (Uhlig et al, 1997). For the 3 inconclusive trials there was reasonable confidence in the findings for one trial (Egger et al, 1985), and moderate and low confidence for other two (Kaplan et al, 1989; Schmidt et al, 1997), respectively. Across all trials, the levels of confidence were affected by various study design limitations, similar to those in the Group I trials, including: use of unstructured non-validated rating systems (Egger et al, 1992), non-blinded study personnel responsible for behavior ratings (Uhlig et al, 1997) or effectiveness of blinding not verified or questionable (Egger et al, 1985; Egger et al, 1992; Boris et al, 1994; Schmidt et al, 1997), use of single source for behavioral outcome measures not confirmed by other sources of measurement (Egger et al, 1992; Boris et al, 1994; Uhlig et al, 1997), all/part of data not statistically analyzed (Schmidt et al, 1997) or incomplete presentation of data (Schmidt et al, 1997).

Collectively, these trials provided reasonably acceptable evidence (Carter et al, 1994; Egger et al, 1985) or suggestive findings (Boris et al, 1994; Egger et al, 1992; Kaplan et al, 1989; Schmidt et al, 1997; Uhlig et al, 1997) to conclude that certain children with ADHD or other problem behaviors may exhibit a unique intolerance to a variety of foods and food components, including artificial colors. While a general increase in hyperactive behaviors in ADHD/behavior problem children has been associated with food intolerance (Boris et al, 1994; Egger et al, 1985), there is some evidence that exposure of such children to various individual provoking food items may

specifically result in small but adverse behaviors associated more with irritability, fidgetiness, sleep problems, and impulsiveness, than attention deficit and learning deficiency which are related to the hyperactivity syndromes (Carter et al, 1993). Carter et al (1993) also suggested that subsequent investigations should consider the extent to which synergistic or additive effects may occur from exposure to multiple provoking food items. The 1982 NIH Consensus Conference report had also identified the need for further research on the synergistic effects of dietary components (NIH, 1982). While it has been suggested that food intolerance in ADHD children may involve some type of immunologic process, there are conflicting results regarding atopy and IgE (Boris et al, 1994; Egger et al, 1985; Egger et al, 1992; Schmidt et al, 1997), but the desensitization trial by Egger et al (1992) attributed desensitization induced tolerance to provoking foods as an immunological response most likely involving a non-IgE cellular response to antigen rather than an antibody mediated immunization**.**

V Overview of Animal/Laboratory Studies

A variety of animal laboratory studies have been conducted to determine whether there is any biological support for a possible links between artificial food colors and Hyperactivity. A fairly representative group of these studies are referenced as supportive information in this review (see attached *References: IIa. Animal/Laboratory Studies*). Early studies reported that the food color, erythrosine (Red No. 3), was shown *in vitro* to inhibit the uptake of neurotransmitters, specifically including dopamine, by nerve cells (Lafferman and Silbergeld, 1979; Logan and Swanson, 1979). However, subsequent experimental information found that effect was due to nonspecific interactions of erythrosine with many biological membranes rather than a specific neuronal effect, which made a link with Hyperactivity very difficult to discern (Mailman and Lewis, 1983). To help determine whether behavior may be affected by erythrosine, investigators also began conducting animal studies but these produced rather variable results, providing no clear evidence that erythrosine had any significant adverse effects on behavioral functions. Some of those early animal studies showed no behavioral effects of erythrosine in either developing or adult animals (Goldenring et al, 1981; Mailman et al, 1980), while others did report positive effects but often with no clear dose response or at high dose levels (see review in background bibliography by Silbergeld and Anderson, 1982). Subsequently, additional laboratory studies have been conducted related to erythrosine and other color additives. Among the 4 additional erythrosine studies, reviewed in this report, one showed that erythrosine does not appreciably penetrate the blood brain barrier in adult rats (developing animals not investigated) (Levitan et al, 1985); another showed that erythrosine does not affect activity in adult mice in the dark or under irradiated light (Galloway et al, 1986); and two developmental neurotoxicity studies, one with rats and one with mice, showed that there was no evidence of neurobehavioral toxicity in developing rats from dietary exposure to erythrosine up to the highest dietary level used of 1% (Vorhees et al, 1983) and only few minor behavioral milestone changes in male mice but only at the highest dietary level of 0.045% erythrosine (Tanaka, 2001). Based on the above animal/laboratory information, there appears to be no convincing evidence that can be extrapolated as being supportive of a link specifically between erythrosine and Hyperactivity in children.

The remaining animal/laboratory studies related to foods colors that were reviewed in this report

provided a variety of additional interesting information. *Two developmental neurotoxicity studies, one using rats and the other mice, assessed the effects of dietary exposure to Red 40 (allura red AC). The study with rats, using the higher dietary levels of 2.5 to 10%, found that all dietary levels of Red 40 produced both physical and behavioral toxicity in the developing rats (Vorhees et al, 1983). The study with mice, using lower dietary levels or Red-40 at 0.42% to 1.68%, found no effects on behavioral development at any dietary level and only limited effects on maze performance, of questionable relevance, only at the highest dietary level (Tanaka, 1994). *A developmental neurotoxicity study of amaranth (Red No. 2) was conducted in mice using dietary levels from 0.03% to 0.27% (the 0.03% dietary level is equivalent to a dose level of 50 mg/kg/day). All dietary levels of amaranth significantly affected several measures of behavioral development but no effects on maze performance and inconsistent changes in activity (Tanaka, 1992). *A single generation toxicity study with rats exposed *in utero* showed carmoisine at dose levels up to 1200 mg/kg/day delivered in the diet showed no overt behavioral effects but some general signs of toxicity, such as decreased body weights, starting somewhat at 400 mg/kg/day (Ford et al, 1987). *A combined reproductive and developmental neurobehavioral toxicity study was conducted in mice with dietary tartrazine (Yellow 5) at levels of 0.05 to 0.45%. Only the highest dietary level of tartrazine (0.45%) produced significant adverse effects on a few indices of behavioral development in developing mice (Tanaka, 2006). *A behavioral development study showed that postnatal injections of sulfanilic acid, a metabolite of azo dyes such as Yellow 5 (tartrazine) and Yellow 6, produce several significant behavioral changes in developing rats that are dissimilar to those effects of 6OHDA injection, which is considered an animal model of Hyperactivity, but these findings have little relevance to humans because of ADME differences (Goldenring et al, 1982) *A reproductive and neurobehavioral development study of *lac* dye, a natural color additive, was conducted in mice at dietary levels of 0.15 to 0.6%. There were variable statistically significant effects on behavioral development and function across all dietary levels in both sexes, most consistently at the highest dietary level of 0.6%, also occurring at the 0.3% dietary level, and occasionally at the lowest 0.15% dietary level; body weights of both sexes were significantly decreased toward the end of lactation with females noted as more affected than males (Tanaka, 1997). *A very interesting *in vitro* study with neuroblastoma cell cultures showed a potential *synergistic* neurotoxicity in inhibiting neurite outgrowth (an *in vitro* model of neuronal cell differentiation) with two combinations of food additives, specifically Brilliant Blue + MSG (l-glutamic acid) and Quinoline Yellow + aspartame (Lau et al, 2006).

In 1982, the NIH Consensus Conference (NIH, 1982) concluded that there was a need for epidemiological studies to include addressing possible genetic, developmental and environmental factors which may be causal and serve as predictors of effect, and animal studies to obtain basic relevant biological information. Thus far, based on the representative studies identified above, the primary contribution of the animal/laboratory studies in providing biological information linking artificial food colors and ADHD seems to have focused on identifying particular color additives that may have a potential for causing behavioral (neurotoxic) effects. Even in this regard additional laboratory testing in needed to better characterize neurotoxic potential of these chemicals, particularly at lower levels of exposure, and to determine whether sensitivity to their potential neurotoxicity may be modulated by genetic polymorphisms or by some synergistic interaction with other chemical substances. Hopefully, this type of information can at least help clinical investigators to prioritize color additives of interest and focus future clinical studies on the more suspect chemicals. However, laboratory investigations need to devote increasingly more attention to systematically exploring the possible biological processes that may underlie

links between food colors and ADHD or other related behavioral disorders of childhood and, as necessary, develop innovative new experimental approaches for chemical testing, for example developing an *in vivo* protocol to complement the *in vitro* demonstration of possible *synergistic* neurotoxicity (Lau et al, 2006).

VI Possible Biological Processes/Mechanisms

Attempts to identify the biological process(es) underlying the proposed relationship between artificial food colors and problem behaviors in children, such as ADHD, is complicated by the multitude of possible scenarios due to the broad array of basic questions about the nature of this proposed relationship. For example, are the colors acting through some toxic, physiologic, allergic or other immunologic process? Are the major behavioral effects caused by one particular color, by the combined action of multiple colors, or by some interaction, perhaps synergistic, with other component(s) in the food? Are these color effects associated with some factor(s) that predispose children to ADHD or other types of behavioral pathology, or could the color effects be associated with some predisposing factor(s) not necessarily related to behavioral disorders? Although many investigators have speculated about these various issues, most of these basic questions still remain largely unanswered.

The considerations of possible biological processes that may underlie a relationship between artificial food colors and problem behaviors, such as ADHD, in children have can be summarized into several broad categories:

Neurochemical

As noted previously, one of the earliest proposed biological mechanisms linking food color additives and hyperactivity involved a defective neurochemical process affecting synaptic availability of certain neurotransmitters, particularly dopamine. The basis for this original proposal was eventually considered inconclusive. However, a dopaminergic or other neurotransmitter involvement is still considered likely based on the view that altered dopaminergic neurotransmission may be involved in the pathophysiology of ADHD (Brookes et al, 2006; Sonuga-Barke, 2003). Since therapeutic treatments are known to positively modulate the dopaminergic system (Banergee et al, 2007), it seems logical that treatments (colors) that can trigger or exaggerate ADHD behaviors may possibly be expected to negatively modulate the dopaminergic system. Hopefully, investigators will take full advantage of the recent identification of several gene variants associated with susceptibility to ADHD that include dopamine receptor and dopamine transporter genes (Banerjee et al, 2007; Farone et al., 2001; Farone et al, 2005) to pursue productive investigations of the role of dopamine in the effects of colors on ADHD.

Many environmental factors can increase histamine release, including infections as well as many food items and certain artificial food colors. This, together with the frequent claim that food allergy/intolerance is a cause of hyperactivity has led to the suggestion that the current focus on a dopaminergic mechanism in ADHD needs to be extended to histamine (Stevenson et al, 2007a). Since genetic polymorphisms involving the histamine N-methyltransferase gene can impair histamine clearance and the histamine (H3) receptors are present in the brain, this provides a possible mechanistic basis for gene-food interactions associated with ADHD. Indeed, some tentative information that genetic variants related to histamine may modulate behavioral responses to artificial colors in

some children was suggested in a study report by Stevenson et al (2007).

Genetic Processes

The possibility that genetic processes may underlie the link between colors and ADHD stems from the fact that there is a strong genetic component for ADHD (Banerjee et al, 2007; Goodman and Stevenson, 1989; Stevenson, 2006). In addition to the genetic component, there appears to be a variety of interacting biological and environmental factors that may be associated with expression of the ADHD (Banarjee et al, 2007) and food may be one of those risk factors that may elicit or exaggerate, but not cause, hyperactive behaviors in some children (Cruz and Bahna, 2006; Mattes, 1983; NIH, 1982; Schab and Trinh, 2004; Wender, 1986). But it remains to be determined whether the genetic variants associated with ADHD may also modulate sensitivity to food additives or, generally, the development or expression of food intolerances. Some suggestion for this does come from results presented by Rowe and Rowe (1994) in which they reported that more hyperactive children reacted to color challenge than normal children, suggesting that the certain genetic elements associated with predisposing children to hyperactivity may also predispose some of those children to a sensitivity to food colors. Also, as noted earlier, there is some tentative information that genetic variants related to histamine, which itself may be associated with ADHD, may modulate behavioral responses to artificial colors in some children (Stevenson et al, 2007). However, as Schon and Trinh (2004) point out, the possible contribution of artificial colors in triggering the expression of ADHD must contend with the incongruity that the pattern of behaviors reported by Rowe and Rowe (1994) following the color challenge (although tartrazine only) differ from the behaviors associated with ADHD. A similar incongruity exists for food intolerance which is also reported to elicit behaviors that are not characteristic of ADHD (Carter et al, 1993). Additional systematic experimental studies are needed to provide more systematic information in this area.

Food Intolerance/Allergy/Immunologic

The adverse effects of artificial food color, although limited and affecting only a small group of children with problem behaviors, such as ADHD, have not been consistently associated with atopy and are now generally thought not to be caused by an allergic reaction mediated through an IgE mechanism (Bateman et al, 2004; MacGibbon, 1983; Pollock and Warner, 1990). The color effects are more likely to occur through some pharmacologic effects such as a non-IgE dependent histamine release (Bateman et al, 2004).

An observation was made that most of the children who are anecdotally reported to improve on the Feingold diet do not show adverse behavioral response to the controlled challenge with color additives which indicates that other factors in the diet, not the artificial food colors, are the key dietary variables for those children (Bishop, 1983; Stare et al, 1980). Investigators began broadening Feingold's original hypothesis to restrict not only food colors and flavors, but also any food items that were assumed or suspected of causing an adverse reaction and reported findings suggesting that multiple food items can provoke adverse behavioral reaction (Kaplan et al, 1989; Schmidt et al, 1997). Other investigators used challenge trials which suggested that some ADHD children have intolerance to a variety of food items, not limited to colors or additives, which may cause or exaggerate adverse behaviors. In an interesting food desensitization study, ADHD

children were desensitized which produced in these children a tolerance to food items that previously provoked adverse behavioral reactions (Eggers et al, 1992). The authors attributed this to an immunological response most likely involving a non-IgE cellular response to antigen rather than an antibody mediated immunization**.**

It should be noted that, although understanding these modes of action (neurochemical and genetic processes) will aid in filling the data gaps, technical limitations presently exist in connecting basic animal neurochemical anatomy/physiology or genetic/epigenetic factors with complex and subtle human behavioral characteristics.

VII Overall Evaluation & Interpretation of Available Information

The interpretation of and conclusions drawn from the studies reviewed in this report includes consideration of the differing dietary conditions (defined diet, food elimination diet), challenge items (diet, artificial color(s), provoking food items), and study population (ADHD, heterogeneous problem behaviors, general population). Overall interpretation of the significance of the reported findings is complicated by the methodological limitations in many of these studies affecting the level of confidence in the data and the occurrence of inconclusive or inconsistent findings, which in several trials consisted of small treatment effects for subjects many of whom were selected for their reported diet responsiveness prior to the study.

Group I Trials: Assessment

The focus of the 26 Group I trials was to assess the validity of Feingold's hypothesis relating food additives and ADHD and other problem behaviors in susceptible children. The clinical trials were typically designed as double blind, placebo controlled studies either to assess the effectiveness of the Feingold diet as a dietary treatment of ADHD or to test the role specifically of artificial colors in causing or provoking adverse behaviors related to ADHD or other problem behaviors in children. Various caveats in the study designs and uncertainties in the study results make interpretation of the study findings difficult and limited in scope.

Pre-1982 Clinical Trials: Collectively, across the 16 pre-1982 trials, the numbers of trials reporting findings of either improved behavior on Feingold's diet or adverse reactions to color challenge were mixed: 2 positive, 8 inconsistent, and 6 negative. Based on an NIH Consensus Committee's review in 1982 of the extant information (including 16 trials conducted prior to 1982 and reviewed in this project), the committee report alluded to differences and inadequacies in the designs of the controlled clinical trials making analysis difficult, but concluded that the available studies "*did indicate a limited positive association between 'the defined diets' and a decrease in hyperactivity...that involved only a small proportion of patients; furthermore, the decreases in hyperactivity were not observed consistently. Studies indicated that some hyperactive children on a defined diet experienced an increase in hyperactivity when given moderate doses of artificial food colors.....by only a small group of patients and the increase was not consistently reported by teachers, parents, and other observers*". This highly qualified conclusion fairly represents this reviewer's assessment of the information from the Group I clinical trials conducted prior to 1982.

Post-1982 Clinical Trials: Collectively, across the 10 post-1982 trials, the numbers of

trials reporting findings of adverse reactions to color challenge were mixed: 3 positive, 3 inconsistent, and 4 negative. Among the 3 positive trials associating adverse behavioral effects with artificial color challenge, there was reasonable confidence in only one study showing clear dose response behavioral effects at each challenge dose of tartrazine from 1 to 50 mg primarily in problem behavior children but also in several children without behavioral problems (Rowe and Rowe, 1994), moderate confidence in one trial reporting 2 children (one hyperactive, the other non-hyperactive with problem behaviors) responding to 50 mg tartrazine or carmoisine with adverse behaviors that may or may not be characteristic of hyperactivity (Rowe, 1988), and low confidence in the third trial reporting several children responding to intermittent challenge with 10 mg tartrazine and sunset yellow (Sarantinos et al, 1990). There were mixed levels of confidence in the remaining trials reporting inconsistent findings (*reasonable* -Thorley, 1984; *moderate*-Salamy et al, 1982; *low*- David, 1987; *low*-Wilson and Scott, 1989) and negative findings (*reasonable*- Pollock and Warner, 1990; *moderate*- McCann et al, 2007; *limited*-Bateman et al, 2004). Particular note should be made of the dose response trial with tartrazine, conducted by Rowe and Rowe (1994) was considered to present some of the more reliable and significant findings. The clear dose related behavioral changes in response to tartrazine challenge across dosages of 1 to 50 mg did show evidence of adverse general behavioral effects primarily in suspect hyperactive children with problem behaviors (and also several control children). The behavioral effects elicited by the tartrazine challenges, however, involved irritability, fidgetiness and sleep problems which are not typically representative of hyperactivity related behaviors. Several other investigators also reported behavioral responses to color challenge that were not particularly characteristic of ADHD (Mattes and Gittelman, 1978; Rowe, 1988; Sarantinos et al, 1990); the citations reporting non-ADHD types of behavioral effects are highlighted in Table 1. Carter et al (1993) also reported similar types of non-hyperactive behavioral responses to provoking food challenges in their study group of ADHD children. The study by Rowe and Rowe (1994) is also notable for reporting that more of the children assessed hyperactive reacted to tartrazine color challenge than normal children. Given the fact that there is a strong genetic component to hyperactivity, more hyperactive than control subjects responding to the adverse behavioral effects of tartrazine suggests that the genetic elements predisposing children to hyperactivity may also predispose some of those children to sensitivity to tartrazine.

The results from all 26 Group I controlled clinical trials, overall, suggest the conclusion that certain subgroups of children with problem behaviors that may or may not be related to ADHD/ADD and, possibly, certain children from the general population without particular behavioral problems may exhibit a unique intolerance to artificial food colors resulting in typically small to moderate adverse behavioral changes which may not necessarily be characteristic of the ADHD syndromes.

Group II Trials: Assessment

Collectively, these trials provided reasonably acceptable evidence in 2 trials (Carter et al, 1994; Egger et al, 1985) or suggestive findings in 5 trials (Boris et al, 1994; Egger et al, 1992; Kaplan et al, 1989; Schmidt et al, 1997; Uhlig et al, 1997) to conclude that certain children with ADHD or other problem behaviors may exhibit a unique intolerance to a

variety of foods and food components, including artificial colors. While a general increase in hyperactive behaviors in ADHD/behavior problem children has been associated with food intolerance (Boris et al, 1994; Egger et al, 1985), there is some evidence that exposure of such children to various individual provoking food items may specifically result in small but adverse behaviors associated more with irritability, fidgetiness, sleep problems, and impulsiveness, than attention deficit and learning deficiency which are related to the hyperactivity syndromes (Carter et al, 1993). Carter et al (1993) also suggested that subsequent investigations should consider the extent to which synergistic or additive effects may occur from exposure to multiple provoking food items. The 1982 NIH Consensus Conference report had also identified the need for further research on the synergistic effects of dietary components (NIH, 1982). While it has been suggested that food intolerance in ADHD children may involve some type of immunologic process, there are conflicting results regarding atopy and IgE (Boris et al, 1994; Egger et al, 1985; Egger et al, 1992; Schmidt et al, 1997), but the desensitization trial by Egger et al (1992) attributed desensitization induced tolerance to provoking foods as an immunological response most likely involving a non-IgE cellular response to antigen rather than an antibody mediated immunization.

One particularly confusing and contentious issue that has been raised by several clinical investigators and reviewers requires some attention primarily because it casts some doubt on the results from a number of clinical trials investigating artificial colors. The issue is whether the approximately 27 - 30 mg of mixed artificial colors, used in a number of clinical trials, represents a level of artificial color too low to enable detection of any treatment related effects (King, 1984; Rapp, 1982; Rimland, 1983; Rippere, 1983; Schab and Trinh, 2004). Rapp (1982) voiced the opinion that enough food coloring must be used in challenge to produce symptoms and further that the quantity should be tailored to the amount needed to cause an individual child to have symptoms, this to be determined prior to the challenge study. This latter suggestion was actually considered in two studies conducted by Mattes and Gittelman (1978, 1981). In the earlier study (Mattes and Gittelman,1978) the challenge dosage of artificial color was based on an initial dose-ranging trial with the test subjects and in the later study (Mattes and Gittelman,1981) incremental daily dosing up to 78 mg/day was used, and in both studies multiple outcome measures detected no significant color challenge effects (in the latter study parent ratings indicated that several children were responding to the color challenge but this was not supported by multiple other subjective and objective behavioral measures). There are several other observations that might suggest that levels of artificial color, within reason, may not necessarily be an important determinant of behavioral effects in studies assessing the association between food colors and problem behaviors such as hyperactivity of childhood. Several other observations are relevant. First, clinical challenge studies assessing behavioral effects using high doses of mixed artificial colors (Swanson and Kinsbourne, 1980; Mattes and Gittelman, 1981; Thorley, 1984; David, 1987; Rowe, 1988; Pollock and Warner, 1990) have generally shown no more dramatic, reliable or conclusive behavioral effects than studies using lower or moderate dose levels (Harley et al, 1978; Goyette et al, 1978; Rose 1978; Levy and Hobbes, 1978; Levy et al, 1978, Mattes and Gittelman, 1978;Williams et al, 1978, Conners, 1980; Conners et al, 1980; Weiss et al, 1980; Wilson and Scott, 1989; Sarantinos et al, 1990; Bateman et al, 2004; McCann et al, 2007). The only evidence of a dose response effect for any color was reported by Rowe and Rowe (1994) in which tartrazine reportedly affected behavior across all dosages from 1 to 50 mg with a ceiling effect (magnitude but not duration) above 10 mg. Second, in many clinical color challenge studies that recruit children who are responsive to the Feingold diet, parents commonly

report that their children exhibit rapid and dramatic deterioration of behavior with even minor infractions of the Feingold diet (Adams, 1981; Conners, 1980; Mattes, 1983; Mattes and Gittelman, 1978; Pollock and Warner, 1990; Rowe, 1988). Feingold even asserted that minute amounts of foods containing synthetic flavors, colors and salicylates is sufficient to cause a recurrence of the hyperactivity behavioral pattern within several hours which persist for up to several days (Feingold, 1975; Ribon and Johsi, 1982). These observations clearly suggest that dosage levels of artificial color, within reason, may not necessarily be an important determinant of behavioral effects in clinical trials assessing the association between food colors and problem behaviors such as hyperactivity of childhood.

Overall, the available information from all 33 trials does not support a causal relationship of either food intolerance in general or artificial food colors/preservatives in particular with ADHD or other problem behaviors in children. However, within the context of the caveats associated with these studies the findings do suggest that small subpopulations of susceptible children with ADHD and/or other problem behaviors, and possibly susceptible children from the general population, may exhibit similar behavioral reactions, not necessarily related to ADHD, to a variety of foods and food ingredients, not limited to artificial food colors and preservatives. The pattern of study findings, in particular the similarity in behavioral response to a wide variety of foods and food ingredients, including colors, suggests that these effects are not the result of inherently behavioral neurotoxic properties of the food chemicals but that the behavioral changes result from some unique food intolerance to which certain individual children may be predisposed, possibly associated with some genetic or epigenetic factors.
Consequently, a parsimonious assessment of the available information concludes that small subpopulations of susceptible children with ADHD or other problem behaviors and, possibly, certain susceptible children from the general population may be predisposed to a unique intolerance to a variety of foods and food ingredients, not limited to artificial food colors and preservatives, that may be associated with adverse behavioral responses, including non-hyperactive behaviors such as irritability, restlessness and sleep disturbances, and physical responses.

VIII Identification of Information Gaps/Suggested Additional Study

- Information which might throw some light on the mechanism(s) of production of adverse reactions could be obtained by studying differences in the genetic, immunological and pharmacological background of reactors and of non-reactors (GI factors would be of importance, too). Without this understanding of what happens in sensitive human subjects, it may not prove possible to develop an animals or *in vitro* models for predictive testing of new additives (MacGibbon, 1983). Additional clinical studies are needed not only to confirm/determine that adverse behavioral responses are elicited by artificial colors and intolerant foods in hyperactive and normal (non-hyperactive) children but also to compare the nature and extent of these responses between these two groups of children and to understand the mechanism underlying these behavioral responses. Do the responses between normal and hyperactive children differ in nature and/or severity? Are they exhibiting the same types of behavioral changes but of different magnitudes/severities? Are responses to colors and food items categorically different or are they similar enough to suggest both

represent food intolerance? Is there any identifiable genetic or polymorphic predisposition to the adverse behavioral responses? Are these behavioral changes due to some direct effect on the nervous system or are they secondary to some immunologic, allergenic, or other physiological effects of treatment? Subsequent studies should compare the behavioral responses of hyperactive children and of normal (non-hyperactive) children to artificial colors and to provoking food items. Specific physical and immunological assessments should be included in the study design and genetic analyses conducted.

- There are known genetic variants associated with ADHD (Faraone et al, 2001; Faraone et al, 2005; Goodman and Stevenson, 1989). But it is not known whether these proposed genetic variants may also modulate sensitivity to food additives or may be associated with the development or expression of food intolerances. Clinical experimental studies to provide this information should be conducted. Some tentative information that genetic variants related to histamine may modulate behavioral responses of children was suggested in a study report by Stevenson et al (2007). Species differences in toxicity, for example the significantly different neurotoxic potential of benzaldehyde between rats and mice (Kluwe et al, 1983), indicate genetic based differences in susceptibility. Can this be demonstrated for behavioral effects of other food chemicals? Correlative information should be developed in laboratory studies to determine whether genetic modulation of behavioral sensitivity to food additives/colors can be demonstrated in animal models.

- In the food intolerance study conducted by Carter et al (1993), various treatment related effects from challenge with various provoking food items were found, but the small magnitude of these effects make their biological relevance unclear. Other studies have reported similar findings of small treatment related (color challenge and provoking food challenge) adverse behavioral effects. Possibly, combined exposure to multiple types of provoking food items may be necessary to elicit an additive effect of a greater magnitude behavioral response, as suggested by Carter et al (1993). Although the diet crossover study (multiple provoking food items) by Kaplan et al (1989) was inconclusive, their findings did suggest appreciable effects in some of the children. A recent *in vitro* laboratory study has reported possible synergistic effects of several food additives (Lau et al, 2006). There is little systematic information about combination effects of chemical substances, particularly with regard to potential adverse behavioral effects. The possibility of synergistic or additive adverse behavioral effects from exposure to multiple provoking food items should be addressed in subsequent studies.

Considerations in Study Design

The design and conduct of any clinical trial to assess the relationship between artificial colors and problem behaviors, such as ADHD, in children must adhere to basic principles of experimental design and study conduct. The following represent the minimum issues that should be considered to maximize confidence in the reliability of the study findings:

- Homogeneity and characterization of sample
- Randomization to treatment
- Crossover designs with subjects serving as own control
- Counterbalanced treatment/challenge order

- Double-blind/placebo-controlled challenges
- Placebo and challenge indistinguishable
- Verification of effectiveness of blinding, particularly those rating behaviors
- Age-appropriate outcome measures
- Use of validated measures (detect behavior differences/sensitive to treatment effects)
- Individualized/target behavior checklists used by all raters in study
- Use of individualized ratings with other standard/validated ratings
- Multiple sources of outcome measures (e.g., parents, teachers, objective tests, etc.)
- Analysis of all data using appropriate statistical procedures
- All data should be presented at least in summary form (tables/graphs)
- Data interpretation should consider any inadvertent occurrences during study

IX Conclusion

Exposure to food and food components, including artificial food colors and preservatives, may be associated with adverse behaviors, not necessarily related to hyperactivity, in certain susceptible children with ADHD and other problem behaviors, and possibly in susceptible children from the general population. A parsimonious interpretation of findings from relevant clinical trials indicates that this food related triggering of problem behaviors is not due to an inherent neurotoxic property of the food or food components, including any of the artificial food colors and preservatives, but appears to result from a unique intolerance exhibited by certain predisposed children to a variety of food items and color additives. The etiology of this type of unique intolerance is unclear but may involve genetic or epigenetic factors.

X Comment

As a general observation, there seem to be two possible basic scenarios that could be operative whereby food additive/environmental chemicals would be associated with triggering adverse behaviors such as those related to ADHD or other behavioral disorders of childhood, or even in the general population. One “traditional toxicology” scenario is that certain chemicals may have inherent neurobehavioral toxicity properties which may directly or indirectly (e.g., endocrine or immunologic pathways) affect nervous system function resulting in behavioral deficits. This scenario may be addressed with reliable toxicological testing including adequate neurobehavioral toxicological evaluations as a routine component of the process of chemical safety assessment. The other “non-traditional toxicology” scenario is that the elicitation of problem behaviors by various common foods and food related chemicals may be due not to an inherent neurobehavioral toxic property of these food items and food related chemicals but to some unique hypersensitivity or intolerance in certain children stemming from some genetic/epigenetic/polymorphic related predisposition. This latter scenario of unique hypersensitivity can best be addressed by continuing efforts to understand the biomolecular factors that may predispose an organism to this type of unique disruptive behavioral response to otherwise non-neurotoxic chemical substances.

† The color-coding is for quick visual reference, but it is not required to access and interpret the data.

THE FINDINGS AND CONCLUSIONS IN THIS REPORT HAVE NOT BEEN FORMALLY DISSEMINATED BY FDA AND SHOULD NOT BE CONSTRUED TO REPRESENT ANY AGENCY DETERMINATION OR POLICY.

REVIEWS and CRITIQUES
33 CLINICAL TRIALS RELATED TO ARTIFICIAL FOOD COLORS and ADHD in CHILDHOOD and RELATED PROBLEM BEHAVIORS

Thomas J. Sobotka, Ph.D.

June 13, 2010

(Updated August 23, 2010)

(Chronological Order)

Conners CK, Goyette CH, Southwick DA, lees JM, and Andrulonis PA. Food additives and hyperkinesis: A controlled double-blind experiment. Pediatrics, 58 (August, No.2), 1976: 154-166.

Harley JP, Ray RS, Tomasi L, Eichman PL, Matthews CG, Chun R, Cleeland CS and Traisman E. Hyperkinesis and food additives: testing the Feingold hypothesis. Pediatrics. 1978, Jun; 61(6):818-28. (I)

Goyette CH, Connors CK, Petti TA and Curtis LE. Effects of Artificial Colors on Hyperkinetic Children: A Double-Blind Challenge Study. Psychopharmacology Bulletin, 14 (No.2), 1978: 39-40.

Harley P, Matthews CG and Eichman P. Synthetic Food Colors and Hyperactivity in Children: Double-Blind Challenge Experiment. Pediatrics, 62, 1978: 975-983. (II)

Levy F and Hobbes G. Hyperkinesis and Diet: A Replication Study. American Journal of Psychiatry, 135 (No.12), 1978: 1559-1560.

Levy F, Dumbrell S, Hobbes G, Ryan M, Wilton N and Woodhill JM. Hyperkinesis and Diet: A Double-Blind Crossover Trial With A Tartrazine Challenge. Medical Journal of Australia, 1 (No.2, January), 1978: 61-64.

Mattes J and Gittelman-Klein R. A Crossover Study of Artificial Food Colorings in a Hyperkinetic Child. American Journal of Psychiatry, 1978, 135 (No.8): 987-988.

Rose TL. The Functional Relationship Between Artificial Food Colors and Hyperactivity. Journal of Applied Behavior Analysis, 115 (No.4, Winter), 1978: 439-446.

Williams JI, Cram DM, Tausig FT and Webster E. Relative Effects of Drugs and Diet on Hyperactive Behaviors: An Experimental Study. Pediatrics, 1978, 61: 811-817.

Conners CK, Goyette CH and Newman EB. Dose-Time Effect of Artificial Colors in Hyperactive Children. Journal of Learning Disabilities, 13 (No.9, November), 1980: 48-52.

Conners CK (a). Artificial colors in the diet and disruptive behavior. In R Knights and D Bakker (Eds.) *Treatment of Hyperactive and Learning Disordered Children: Current Research*. Baltimore, University Park Press, 1980: 113-120.

Swanson JM and Kinsbourne M. Food dyes impair performance of hyperactive children on a laboratory learning test. Science, 1980, 207 (March 28): 1485-1487.

Weiss B, Hicks William J, Margen S, Abrams B, Caan B, Citron LJ, Cox C, McKibben J, Ogan D and Schultz S. Behavioral Responses to Artificial Food Colors. Science, 207 (March), 1980: 1487-1489.

Mattes J and Gittelman R. Effects of Artificial Food Colorings In Children With Hyperactive Symptoms: A Critical Review and Results of a Controlled Study. Archives of General Psychiatry, 1981, 38:714-718.

Adams W. Lack of Behavioral Effects From Feingold Diet Violations. Perceptual and Motor Skills, 1981, 52: 307-313.

Spring C, Vermeersch J, Blunden D and Sterling H. Case Studies of Effects of Artificial Food Colors on Hyperactivity. The Journal of Special Education, 1981, 15 (Number 3): 361-372.

Salamy J, Shucard D, Alexander H, Peterson D, Braud L. Physiological changes in hyperactive children following the ingestion of food additives. Int J Neurosci. 1982, May; 16(3-4):241-6.

Thorley G. Pilot study to assess behavioural and cognitive effects of artificial food colours in a group of retarded children. Dev Med Child Neurol. 1984 (February), 26(1):56-61.

Egger J, Carter CM, Graham PJ, Gumley D and Soothill JF. Controlled trial of oligoantigenic treatment in the hyperkinetic syndrome. Lancet. 1985 (Mar 9), 1(8428):540-5

David TJ. Reactions to dietary tartrazine. Archives of Disease in Childhood, 1987, 62: 119-122.

Rowe KS. Synthetic food colourings and 'hyperactivity': a double-blind crossover study. Aust Paediatr J. 1988 (April), 24(2):143-7.

Kaplan BJ, McNicol J, Conte RA and Moghadam HK. Dietary replacement in preschool-aged hyperactive boys. Pediatrics, 1989 (January), 83 (No.1): 7-17.

Wilson N and Scott A. A double-blind assessment of additive intolerance in children using a 12 day challenge period at home. Clinical and Experimental Allergy, 1989, 19: 267-272.

Pollock I, Warner JO. Effects of artificial food colours on childhood behaviour. Arch Dis Child. 1990 (January), 65(1):74-7.

Sarantinos J, Rowe KS and Briggs DR. Synthetic Food Colouring and Behavioral Change In Children With Attention Deficit Disorder: A Double-Blind, Placebo Controlled (Challenge Study). Proc. Nutr. Aust., 1990, 15: 233.

Egger J., Stolla A and McEwen LM. Controlled trial of hyposensitization in children with food-induced hyperkinetic syndrome. The Lancet, 1992 (May), 339: 1150-1153.

Carter CM, Urbanowicz M, Hemsley R, Mantilla L, Strobel S, Graham PJ, Taylor E. Effects of a few food diet in attention deficit disorder. Arch Dis Child. 1993 (November), 69(5):564-8.

Boris M and Mandel FS. Foods and additives are common causes of the attention deficit hyperactive disorder in children. Ann Allergy. 1994 (May), 72(5):462-8.

Rowe KS, Rowe KJ. Synthetic food coloring and behavior: a dose response effect in a double-blind, placebo-controlled, repeated-measures study. J Pediatr. 1994 (November), 125(5 Pt 1):691-8

Schmidt MH, Mocks P, Lay B, Eisert HG, Fojkar R, Fritz-Sigmund D, Marcus A and Musaeus B. Does oligoantigenic diet influence hyperactive/conduct-disordered children - a controlled trial. European Child & Adolescent Psychiatry, 1997, 6: 88-95.

Uhlig T Merkenschlager A Brandmaier R Egger J. Topographic mapping of brain electrical activity in children with food-induced attention deficit hyperkinetic disorder. Eur J Pediatr, 1997 (July), 156(7):557-61.

Bateman B, Warner JO, Hutchinson E, Dean T, Rowlandson P, Gant C, Grundy J, Fitzgerald C and Stevenson J. The effects of a double blind, placebo controlled artificial food colourings and benzoate preservative challenge on hyperactivity in a general population sample of preschool children. Arch Dis Child. 2004 (June), 89(6): 506-11.

McCann D, Barrett A, Cooper A, Crumpler D, Dalen L, Grimshaw K, Kitchin E, Lok K, Porteous L, Prince E, Sonuga-Barke E, Warner JO, Stevenson J. Food additives and hyperactive behaviour in 3-year-old and 8/9-year-old children in the community: a randomized, double-blinded, placebo-controlled trial. Lancet, 2007 (November 3), 370 (9598):1560-7.

REVIEWS and CRITIQUES

(trials conducted prior to 1982)

(Conners et al, 1976)

Conners CK, Goyette CH, Southwick DA, lees JM, *and* Andrulonis PA. Food additives and hyperkinesis: A controlled double-blind experiment. Pediatrics, 1976 (August), 58 (No.2): 154-166.

In one of the first controlled clinical trials designed to test the effects of the Feingold diet (eliminated artificial colors, flavors and natural salicylates) on the behavior of hyperactive children Conners et al (1976) recruited 15 children aged 6-12 years who were diagnosed with hyperkinetic reaction of childhood and conducted a single cross-over dietary challenge study comparing the Feingold diet with a matched placebo diet. Parents were given lists of approved foods for use in making the appropriate diet for each challenge period but were not told which diet they were preparing (i.e., blinded to treatment). It should be noted, however, that no specific procedure was used to verify the effectiveness of the parents' blinding, for example, parents and children were not asked whether they could differentiate the F/KP and placebo diets. After several weeks of pre-treatment/baseline, each child was given either the Feingold diet (also referred to as the Feingold/Kaiser-Permanente or F/KP diet) or a matched placebo diet for 4 weeks and then crossed over to the other diet for an additional 4 weeks. The order in which the two diets were given (i.e., F/KP first and Placebo second, or Placebo first and F/KP second) was determined randomly for each child. The outcome measures to determine treatment related changes in behavior consisted of : (1) parent and teacher behavioral ratings, using the standard Conners Parent/Teacher Hyperactivity Questionnaire, and (2) a Global Behavior Assessment by the principle investigator based on interviews with parents and teachers and their behavior ratings.

Study Results: The analysis of the whole study group's data showed significantly improved behavior on the F/KP diet relative to control based primarily on teacher behavior ratings with some support from the global behavior assessment but no confirmation from parent behavior ratings which detected no significant overall diet related changes in behavior. However, the teacher's finding of a F/KP diet effect was not consistent due to an inexplicable treatment x order effect in which the F/KP diet improved behavior was seen primarily when the placebo diet was given first and the F/KP diet second but was less apparent when the F/KP diet was given first. Viewing the data for individual subjects revealed that both the teacher and parent behavioral ratings agreed that 4 to 5 of the 15 subjects showed improved behavior on the F/KP diet, but, as noted for the overall group effect, the diet effects for these individual "responders" were also not consistent due to the treatment x order effect.

***Study Assessment and Evaluation*:** Although the authors conclude that the F/KP diet, free of most natural salicylates, artificial flavors and artificial colors, reduces the perceived hyperactivity of some children with hyperkinetic impulse disorder, the findings on which this conclusion is based are inconsistent and equivocal. The small number of subjects, the uncertain effectiveness of the blinding, the absence of objective behavior outcome measures, the discrepancy between teacher and parent behavioral ratings in detecting an overall group effect, and the inexplicable treatment x order effect which mitigated the significance of not only the teacher reported group effect but also the parent and teacher agreement in identifying 4 to 5 individual children who responded positively to the F/KP diet, lessen the confidence in the reliability and biological significance of the study's findings. The reviewer concludes that this

study provides no credible evidence of any effects of the Feingold diet on childhood hyperactivity.

(Harley et al, 1978: Phase I)

Harley JP, Ray RS, Tomasi L, Eichman PL, Matthews CG, Chun R, Cleeland CS and Traisman E. Hyperkinesis and food additives: testing the Feingold hypothesis. Pediatrics, 1978 (June), 61(6):818-28 (I)

In an effort to test the behavioral effects of the Feingold diet on childhood hyperactivity using objective laboratory and classroom outcome measures, in conjunction with subjective parent-teacher ratings, Harley et al (1978) recruited 46 male hyperactive children (36 school aged at 6-13 years and 10 pre-school aged at 3-6 years) to conduct a double-blind dietary crossover study comparing the Feingold diet with a matched control diet. Throughout the study all of the food (experimental F/KP diet and control diet) for each family was provided weekly by the investigators. Parents, teachers and study personnel were blinded to treatment condition. Efforts were made to verify blinding effectiveness at the end of the study by asking parents if they could identify when they were on the F/KP experimental diet; none of the parents were able to respond correctly. Procedurally, after an initial 2 week baseline on normal diet, the children were randomly assigned to one of two diet conditions (control or F/KP). They were given their initially assigned diet for 3 - 4 weeks (the study was run in three phases and duration of diet exposure varied across phases), then crossed over to the other diet for an additional 3 - 4 weeks. In addition to a neurological and physical examinations, the primary behavioral outcome measures for school aged subjects included: (1) Conners parent and teacher behavioral ratings, (2) classroom and laboratory/activity observations by trained observers, and (3) neuropsychological evaluations. Observation data and neuropsychological test data were collected for hyperactive subjects and corresponding matched controls. Behavioral measure for pre-school subjects included: parent behavior ratings; laboratory/activity observations; and neuropsychological evaluations (the latter data for preschool subjects was not presented but was discussed).

Study Results: The physical exams of all 46 subjects were normal. Neurologically, 14/36 school aged and 2/10 preschool boys had positive findings. In the analyses of the entire dataset for school aged subjects parent behavioral ratings showed a significant overall F/KP diet effect with improvement in hyperactive behaviors. But this diet effect was not consistently seen due to a significant diet x order effect in which the F/KP related improved behaviors were apparent only when subjects were exposed to the control diet first and the F/KP diet second. The significance of the parent detection of a F/KP diet effect was further mitigated by the fact that teacher behavior ratings did not confirm any diet related changes in group behavior and that, while classroom/laboratory observations and neuropsychological testing discriminated between hyperactive and control children, there was no significant or convincing evidence of an overall effect of the F/KP diet on these behavioral measures. In the analyses of the dataset for preschool subjects parent ratings showed a significant group F/KP diet effect of improved behavior but, in contrast to the school aged subjects, there was no confounding diet x diet order effects. Still, the significance of the parental finding of a F/KP diet effect in preschoolers was mitigated by the lack of confirmation of any diet related effects based on either objective measure of laboratory/activity setting observations or neuropsychological testing (teacher behavior ratings were not collected for preschool subjects). The overall assessment of the dataset for individual school aged subjects did reveal an agreement between parent and teacher behavior ratings

identifying 4 of the 36 children as being responsive to the dietary conditions with behavior improving on the F/KP diet. However, as noted for the overall diet effect based on the parent ratings alone, the diet x diet order effect also mitigated the significance of the F/KP diet effects in these 4 subjects. No mention is made by the authors as to whether the other more objective outcome behavioral measures these 4 children showed any confirmatory diet related changes.
***Study Assessment and Evaluation*:** There are several caveats mitigating the reliability and biological significance of the parental ratings in school aged boys showing a significant F/KP diet related improvement in hyperactive behavior. These caveats include a significant diet x diet order effect and an absence of confirmation with teacher ratings or various objective behavioral test measures (neuropsychological evaluation, classroom observations, and laboratory observations), all of which showed no corresponding F/KP diet related group changes in behavior. There was an agreement between parent and teacher ratings that 4/36 children showed improved behavior on the F/KP diet. However, the significance of this is limited by the confounding diet x diet order effect and the absence of confirmation of any comparable diet related changes based on objective laboratory/activity observations and extensive neuropsychological testing. Also, in the preschool subjects the significance of the parental ratings showing a significant diet effect is also mitigated by the small sample size and the lack of confirmation of a F/KP diet effect based on objective measures of laboratory/activity setting observations and a battery of neuropsychological tests. In view of the questionable reliability and biological significance of the findings, this controlled double-blind dietary cross-over study does not provide any credible support for the contention that the Feingold diet significantly affects disruptive behaviors in either school aged or preschool aged hyperactive children or that the hyperkinetic disorder is triggered or exaggerated by artificial food colors/flavors and salicylates. It is questionable whether these study findings would even support the suggestion that a small subset of hyperactive children may be responsive to the Feingold diet. .

(Goyette et al, 1978)

Goyette CH, Connors CK, Petti TA and Curtis LE. Effects of Artificial Colors on Hyperkinetic Children: A Double-Blind Challenge Study. Psychopharmacology Bulletin, 1978, 14 (No.2): 39-40.

Goyette et al (1978) were among the first clinical investigators to conduct a clinical study involving a double-blind challenge with putative offending agents to address the suggestion that common food additives, specifically artificial colors, may cause or exacerbate behavioral symptoms associated with the hyperactivity disorder of children. Two separate challenge trials were conducted. Trial 1, using a double 2-week crossover design over an 8 week period, tested 16 children (average age of 8.3 years, range 4.7-11.8; sexes not specified) with a behavioral and clinical diagnosis of hyperkinesis and carefully screened prior to the challenge phase to show significant symptom reduction when placed on the F/KP elimination diet under non-blinded conditions. The challenge vehicle was a chocolate cookie with/without a 13.5 mg mixture of all approved artificial food colors (½ the average daily intake). While being maintained on an additive-free diet (not specified in the publication but assumed to be the F/KP diet eliminating artificial colors and flavors and salicylates), each child ingested two cookies per day, one after breakfast and one after lunch (total daily color challenge of 27 mg). Color challenge and placebo cookies were each given in 2-week alternating sequences over an 8 week period. Children were randomly assigned to one of two challenge sequences (ABAB or BABA). Behavioral outcome measures included: (1) parent and teacher behavior ratings (3 times per week using Conners

hyperkinesis index) and (2) a visual motor tracking task conducted 1 to 2 hours after ingestion of color challenge/placebo materials (frequency not specified and not clear whether testing was conducted after the first or second daily cookie). Trial 2 was a basic repeat of Trail 1 with the following differences: (1) a daily parent behavior rating for a specific 3-hour period after ingestion of a single cookie at supper was the only outcome measure obtained; (2) younger children were recruited; and (3) a single 2-week crossover design was used. A total of 13 children were tested (mean age of 6 years, range 3.4-10.2; sexes not specified), including 8 diet-responsive, hyperkinetic children ("*criterion*" group) and 5 children with borderline diet response or borderline hyperactivity ("*borderline*" group). Both Trial 1 and 2 were described as being "double blind" but no description of blinding procedures was given other than use of a placebo and random assignment to challenge sequence. Also, no efforts were described for determining whether any of the study participants were in fact blind to treatment condition.
Trial 1 Results: All subjects were clearly identified in open trial, i.e. non-blind, as responders to the elimination diet (parents rated a 57 percent hyperactive symptom reduction and teachers a 34 percent reduction). However, double-blind challenge with color challenge and placebo revealed no significant challenge effects based on any outcome measure (parent rating, teacher rating, tracking task). The tracking task, however, did suggest a non-significant trend for performance deficits with exposure to the color challenge but not with placebo. Inspection of the individual tracking data suggested that younger subjects tended to show a more noted response to the color challenge and retesting of several younger subjects with color challenge alone suggested that effects tend to occur at 1 hour but not 2 or 3 hours after challenge.
Trial 2 Results: Overall, the 13 subjects in Trial 2 showed a 45 percent mean reduction in behavior problems on the open (non-blinded) F/KP elimination diet (relative change in symptoms between "criterion" and "borderline" children was not specified). A significant color challenge effect was found within 3 hours after ingestion of a single challenge cookie (i.e, ½ average daily intake) without any differences between the color response of the "criterion" and "borderline" groups of subjects and without any challenge x order effect.
***Study Assessment and Evaluation*:** In Trial 1 parent/teacher ratings and a tracking task showed that hyperactive children, who were identified in open trial as definite diet responders, exhibited no significant adverse behavioral response to repeated daily challenge with a mixture of artificial colors at levels equivalent to the average daily adult intake when tested under double-blind placebo controlled challenge conditions. The suggestion of a trend for performance deficit in the tracking task, particularly in the younger children, was not significant. The investigators suggested that the parent and teacher behavioral ratings failed to detect a color challenge effect because those ratings which were conducted 3 times per week may have been insensitive due to the long time span of the observation periods compared to a possible short duration of the color challenge effect. This rationale for dismissing these parent/teacher behavioral ratings as insensitive is inconsistent with the fact that the parent and teacher ratings assess behavioral changes throughout the day; the teacher ratings, in particular, reflect the children's behavior during the periods shortly after ingestion of each cookie at breakfast and lunch. In Trial 2, which used slightly younger subjects than Trial 1, significant group color challenge effects were detected based on parent ratings specifically conducted within a 3-hour period after ingestion of a single cookie at supper. According to the investigators, this particular finding suggests that artificial food color do act to impair and disrupt behavior of hyperkinetic children and firmly establish that artificial colors may be particularly disruptive to younger children. In the absence of any other behavior outcome measure to confirm a treatment related behavioral effect and without a clear description of the blinding procedures and verification of the effectiveness of parent blinding, the reliability of the trial 2 findings of a significant color challenge effect on

behavior based only on parent ratings remains questionable. Also, there is limited support for the conclusion regarding younger children being more sensitive, since in Trial 1 the suggestive trend of a treatment effect particularly in younger children was not statistically significant and in Trial 2 there was no difference in the response to color challenge between the two groups of children one with an average age of 5 and the other slightly older children with an average age of 7 (which is actually close to the average age of 8.3 years for the "older" children in Trial 1). Consequently, overall this study does not provide credible evidence for a significant treatment effect of artificial food colors in F/KP diet sensitive hyperkinetic children. Whatever benefit the children may have appeared to obtain from the open F/KP elimination diet was apparently not attributable to the elimination of artificial food colors alone but possibly to some other food component(s) that were eliminated in the F/KP diet.

(Harley et al, 1978: Phase II)

Harley P, Matthews CG and Eichman P. Synthetic Food Colors and Hyperactivity in Children: Double-Blind Challenge Experiment. Pediatrics, 1978, 62: 975-983. (II)

Harley et al (1978: Phase I) conducted an earlier double-blind dietary crossover study (Phase I) comparing the Feingold/Kaiser-Permanente diet, eliminating artificial colors and flavors and salicylates, with a matched control diet to determine whether the disruptive behaviors of hyperactive children can be improved with the F/KP elimination diet. These same investigators conducted the present Phase II study (Harley et al, 1978: Phase II) designed to select those children who showed the best response to the F/KP diet in Phase I and then to challenge them repeatedly with specified amounts of artificial food colors while being maintained on the F/KP diet. Nine male hyperactive children (mean age 9.3 years), identified as F/KP diet “responders” were selected from the Phase I study (Harley et al, 1978: Phase I). One non-hyperactive control subject was matched with each hyperactive subject on sex, grade, and academic ability. The challenge vehicles were candy bars and cookies with/without a 13.5 mg mixture of all approved artificial food colors (two items were ingested daily throughout the challenge period of the study resulting in a daily color challenge of 27 mg, which the investigators considered the “average” daily intake of artificial colors). A panel of judges determined the color/placebo cookies to be indistinguishable. All study participants were blinded to placebo/challenge code and post-study interviews showed that none of the parents or children could distinguish between the placebo and color challenge items (effectiveness of blinding appeared to be reliably established). Procedurally, the subject and his family were started on the elimination diet during the second 2 weeks of a 4 week baseline period prior to the challenge phase of the study (placement on the elimination diet appeared to be non-blinded). While being maintained on the elimination diet, the challenge phase was conducted over a 9 week period during which subjects were randomly assigned to one of two double cross-over challenge sequences with daily challenges (placebo/2 wks, color/2 wks, placebo/2 wks, color/3 wks or colors 2 wks, Placebo 2 wks, Colors 2 wks, Placebo 3 wks). Behavioral outcome measures included: (1) parent (mother and father) and teacher behavior ratings (2 times per week during baseline and challenge), (2) classroom behavioral ratings by trained observer (2 times per week during baseline and challenge), and (3) neuropsychological testing (at end of baseline and of challenge weeks 2, 4 and 9).
Study Results: During the initial period of baseline with subjects on their ‘regular diets’ (investigators provided no information about composition of the hyperactive children’s regular diet), parent and teacher behavior ratings and classroom observer data showed an expected significant difference between the hyperactive and matched control subjects (hyperactives with

more disruptive behavior). It is significant to note that these same outcome measures found that behaviors of the hyperactive and control subjects were still significantly different during the latter part of baseline when all of the hyperactive subjects and their families were placed openly (non-blinded) on the Feingold elimination diet, indicating the perception by parents, teachers and classroom observers that the Feingold elimination diet had only a nominal, if any, effect on the behavior of the hyperactive children who were selected as being the best diet responders in the earlier Phase I study. Analysis of the group challenge data from parent and teacher ratings, classroom observer data, and neuropsychological testing revealed no significant color challenge effects. Graphic analysis of the behavioral profiles from parent and teacher ratings and classroom observer data for each individual hyperactive subject showed that only 1 of the 9 hyperactive subjects displayed a behavioral profile of parent ratings and classroom observer data that even approximated a significant color challenge effect of increased hyperactive behaviors. Teacher behavior ratings did not appear to support a color challenge effect in that single subject and the neuropsychological test data apparently were not considered in the individual profile graphic analysis.

***Study Assessment and Evaluation*:** The results from this study suggest three conclusions related to the proposed association between artificial color and problem behaviors related to ADHD (hyperactivity of childhood). **(1)** This study provides little support for any substantial or consistent beneficial effects of the Feingold elimination diet on the behavior of hyperactive children. When hyperactive children were assessed while on their 'regular diet', their behavior was, as expected, more disruptive than the control (non-hyperactive) children. But when the hyperactive children were placed openly on the Feingold elimination diet, behavioral assessments found their behaviors remained significantly worse than matched control children. Although the Feingold diet was given under non-blinded conditions, it is significant that parents, teachers and classroom observers perceived the behavior of hyperactive children on the Feingold diet still worse than that of controls, even though these hyperactive children were specifically selected for this study on the basis of being identified as among the best Feingold diet positive responders from an earlier diet study. With the caveat that these observations were made under non-blinded conditions, these results provide no confirmation of any or consistent significant beneficial effect of Feingold's elimination diet on the behavior of hyperactive children. **(2)** This study does not support or confirm the hypothesis that artificial food colors cause or trigger problem behaviors related to hyperactivity of childhood. The overall results of this study based on multiple outcome measures (parent, teacher, classroom observer ratings, and neuropsychological testing) showed no adverse behavioral effects of daily ingestion of a 27 mg mix of artificial colors (considered the average daily intake of artificial colors) by hyperactive children for up to 21 consecutive days and, consequently, do not support or confirm Feingold's hypothesis, based on his clinical observations, (Feingold, 1975) that ingestion by hyperactive children of even minute amounts of foods containing synthetic flavors, colors and salicylates "causes a recurrence of the complete behavioral pattern within two to four hours which persists for one to four days". **(3)** Certain results from this study may be viewed as possibly suggesting that a small subgroup of hyperactive children may be intolerant to artificial food colors and respond with a limited range of behavioral response. In the assessment of behavioral profiles for each individual subject, 1 of 9 subjects exhibited a behavioral profile, based on parent ratings and classroom observer data, which reportedly (data not shown in this paper) suggested disruptive behavior related to color challenge. Although teacher behavior ratings did not support this suggested color challenge effect (and neuropsychological test data were not included in the individual assessments), these results may still be viewed as possibly suggesting that a small subgroup of hyperactive children may be intolerant to artificial food colors and respond with

limited range of behavioral response.

(Levy and Hobbes, 1978)

Levy F and Hobbes G. Hyperkinesis and Diet: A Replication Study. American Journal of Psychiatry, 1978, 135 (No.12): 1559-1560.

A single crossover challenge study of tartrazine was conducted using 7 hyperkinetic children (6 male, 1 female; average 5.2 years old) who were identified as being responsive to an open Feingold/KP elimination diet based on at least a 25% decrease in the Conners behavior ratings conducted by the mothers. (8 subjects were recruited but 1 subject dropped out of the study.) The challenge vehicle consisted of cookies each with/without 1 mg tartrazine (Yellow #5) (four cookies were ingested on each challenge day resulting in tartrazine challenge of 4 mg). Procedurally, all subjects were maintained on a F/KP elimination diet throughout the 4-week challenge period. Subjects were randomly assigned to a placebo or color challenge group and received 4 scheduled cookies daily for 2 weeks. Subjects were then crossed over to the opposite challenge group for a additional 2 weeks of testing. The parents provided the only outcome measures which consisted of a Conners parent behavior rating conducted daily during the 3-hours immediately after challenge and a parent global behavioral assessment at the end of each 2-week challenge period. Other than the use of placebo cookies, no other measures to blind the challenge study were described.

Study Results: In a placebo controlled single crossover challenge study, parent behavior ratings and global assessments revealed no significant adverse effects of daily challenge with 4 mg tartrazine for 2 weeks on behaviors of 7 hyperactive children, identified as responsive to the Feingold/KP elimination diet.

Study Assessment and Evaluation: There are several limitations to this study, including a small sample size of 7 children, no description of efforts to blind the parents to treatment conditions (other than the use of a placebo control), and use of a single source (parents) for outcome measures with no other outcome measures to confirm the parent findings. These caveats notwithstanding, the study provided no evidence that tartrazine at 4 mg per day for 2 weeks has any adverse effect on the behavior of hyperactive children (6 m, 1 f; 5.2 years old) who were reported to be responsive to the Feingold elimination diet.

(Levy et al, 1978)

Levy F, Dumbrell S, Hobbes G, Ryan M, Wilton N and Woodhill JM. Hyperkinesis and Diet: A Double-Blind Crossover Trial With A Tartrazine Challenge. Medical Journal of Australia, 1978 (January), 1 (No.2): 61-64.

Levy et al (1978) used 20 clinically diagnosed hyperkinetic children (19 male, 3 female; 4-8 yrs) to conduct a single crossover challenge study of tartrazine (Yellow #5). (22 children were recruited but 2 children did not complete the challenge portion of the study.) Responsiveness to a Feingold/KP elimination diet was assessed in open trial prior to the challenge portion of this study. The challenge vehicle was a biscuit with/without 1 mg tartrazine (five biscuits were ingested on each challenge day resulting in tartrazine challenge of 5 mg). It should be noted that the "second batch" of color challenge biscuits differed in color from the placebo biscuits, thereby compromising the blind for whatever portion of the study the second batch of color challenge biscuits were used; but it is not clear what portion of the study was actually involved.

Procedurally, following a 2-week pretest baseline period (regular diet), all subjects were place on the F/KP elimination diet over a 4-week period and remained on that diet throughout the rest of the study. After the 4 weeks of diet initiation, half of the subjects were then assigned (random assignment assumed) to a placebo group and half to a color challenge group each receiving their appropriate 5 biscuits daily for a 2-week period. For the next period of 2 weeks, the children were crossed to the opposite treatment group. After completion of the challenge portion of the study, all children were then continued on the F/KP diet for a final 4-week washout period. Multiple outcome measures were taken at the end of each of the test periods described above, including: Conners behavior ratings completed by the mother, teachers and clinicians (psychiatrist and psychologist), and a series of objective psychological response tests to assess activity/motility, attention, impulsivity, perceptual-motor functioning, memory, intelligence and maze performance. It should be noted that the mothers and teachers were asked to rate the child over the 24-hour period before scoring. Blinding efforts included use of placebo biscuits, masking the color difference between placebo and color challenge biscuits, and use of a coded treatment schedule. No efforts to verify effectiveness of the blind were described, for example by asking the parents if they could differentiate the placebo and challenge biscuits.

Test Results: At the end of the first four weeks of open treatment with the F/KP diet, the mothers' behavioral ratings indicated a significant positive diet effect but these ratings were made under non-blind conditions and there was no confirmation of any diet effect based on the teacher or clinician behavior ratings on on a majority of the series of objective psychological tests (the singular exception being a diet related improvement of maze performance, but this effect was not maintained in other study analyses). In the primary analysis of data for the full sample of 20 test subjects, there were no significant adverse effects of daily challenge with 5 mg tartrazine over a 2-week period based on any outcome measure including parent and teacher behavior ratings and any of the series of objective psychological tests. In order to determine whether the effects of diet and of tartrazine challenge might be amplified in the more hyperactive children, post-hoc analysis of the data was conducted for a selected subset of 16 children who were initially rated by teachers as exhibiting higher levels of hyperactivity (i.e. Conners scores $\geq$ 17). This analysis yielded the same general pattern of negative results as that for the full sample indicating that sensitivity to diet and tartrazine challenge was not affected by level of hyperactivity. Additional post-hoc analysis was also carried out to investigate the ability of a select subgroup of mothers (parents) to observe adverse behavioral effects of tartrazine challenge. For this analysis a subgroup of 13 children was selected based on their being suspected as responsive to the F/KP diet, specifically rated by mothers under non-blind conditions as exhibiting a 12 to15% reduction in problem behaviors when placed openly on the diet. Analysis of the mothers' behavior ratings for these children during the placebo/challenge period did indicate a significant adverse tartrazine challenge effect. However, no other confirmatory outcome measure, for example teacher ratings, was included in this particular post-hoc analysis.

Study Assessment and Evaluation: The report of positive improvements in the hyperactive childrens' behavior when placed openly on a F/KP elimination diet is of questionable credibility in view of the fact that, as the investigators point out, it was based solely on the mothers' ratings under non-blinded conditions and was not supported either by teacher ratings or by most of the objective psychological tests. The primary analyses of data for the full sample of 20 hyperactive children showed that daily challenges with 5 mg tartrazine for a 2 week period under placebo controlled conditions did not "trigger" any adverse behavioral effects. Special analysis of the study data for the more hyperactive children showed that the level of hyperactivity does not affect the amplitude of response or sensitivity to the effects of diet or tartrazine challenge.

However, when special analysis of only the mothers' ratings during challenge was conducted for a subgroup of children who reportedly exhibited a 12 to 25% behavioral improvement on the open F/KP elimination diet, there was a significant adverse tartrazine effect. However, the Reviewer considers there to be little, if any, confidence in the reliability or biological relevance of this finding, since the reported diet response of these selected children is based exclusively on mothers' ratings which are not considered credible (see comments above), the preparation of one batch of the challenge biscuits may have compromised the study blinding to some unknown extent, and other confirmatory outcome measures (teacher and clinician ratings and objective tests) were not included in this special analysis. Overall, this study provides no credible evidence of any significant adverse effects of daily exposure to 5 mg tartrazine for 2-weeks in hyperactive children.

(Mattes and Gittelman, 1978)

Mattes J and Gittelman-Klein R. A Crossover Study of Artificial Food Colorings in a Hyperkinetic Child. American Journal of Psychiatry, 1978, 135 (No.8): 987-988.

Mates and Gittelman (1978) designed a placebo controlled double-blind clinical challenge study to maximize the likelihood of demonstrating a food color additive diet effect in one male child (10 years) diagnosed as hyperkinetic and whose parents reported definite behavioral improvement on the Feingold elimination diet and dramatic worsening of behavior, when the diet was violated, with restlessness and irritability lasting from hours to days. The challenge vehicle was a Nutrition Foundation cookie with/without a mix of commonly used artificial colors at a level 1/5 the average daily intake. The cookies, provided by the Nutrition Foundation, were made to be indistinguishable in taste and appearance The study was conducted in two phases both with the child being continued on the Feingold elimination diet: (1) Dose-Ranging Phase – on each of 2 consecutive weeks, the child was randomly determined to receive either placebo or color cookies in progressively greater numbers each week up to six cookies per day; and (2) Multiple Crossover Challenge Phase – a 10-week double-blind multiple crossover challenge in which the child ate 3 cookies a day for two days each week, Wednesday and Thursday. Within each week the same type cookie was eaten on both days and the type of cookie each week was randomized (coin toss). The outcome measures consisted of daily behavior ratings (Conners brief questionnaire) by the mother and teacher daily during the dose-ranging phase and weekly during the multiple crossover phase. Also, at the end of each week, the child, mother and teacher were asked to guess whether that week's cookies were color or placebo.
Study Results: In the dose-ranging phase the mother's ratings showed a behavioral response to the color cookies (but not the placebo) starting after two cookies and a more exaggerated response after five cookies. The mother's ratings went from a baseline value of zero to a score of 15 (the usual cutoff score for hyperactivity) but restlessness and irritability were the predominant behaviors noted. The teacher's ratings, however, did not show any behavioral changes. Based on the mother's ratings during dose-ranging, three cookies per day were selected as adequate for the multiple crossover challenge phase of the study. In the multiple crossover challenge phase, neither the mother's nor the teacher's behavior ratings detected any adverse effects of challenge. The teacher and child were equally unable to guess reliably the type of cookie used each week (5/10 correct for both), but the mother was able to guess correctly 8/10 weeks (in guessing, the mother focused on irritability and fidgetiness as the main effects of the active/color cookies).
Study Assessment and Evaluation: The reviewer concurs with the authors' conclusion that the

results of this study do not support the contention that artificial food colors are instrumental in inducing significant changes in hyperactive symptomatology in a diagnosed hyperkinetic child even with parental anecdotal reports of marked positive response to the Feingold diet and dramatic behavioral deterioration with diet violation. The reviewer agrees with the investigators' view that the severe reaction reported by the mother, but not the teacher, during the Dose-Ranging phase was apparently not related to the active/color cookies, since this reaction did not occur with the same dosage during the Multiple Crossover phase. The investigators suggested that, since the mothers could correctly guess the type of cookie being used each week during the Multiple Crossover phase of this study, some aspects of the child's behavior may have been affected by the color challenge cookies but that this effect appeared to involve types of behavior, such as irritability, that are not typically associated with hyperkinesis and apparently not detected with the typical rating scales used for hyperactive children. This rationale of the authors is not consistent with the fact that the mother's behavioral ratings, but not the teacher's, did detect behavior responses in the Dose-Ranging phase of this study including irritability and restlessness (fidgetiness) (*see note below). Alternatively, it is possible that the mother was able to correctly guess which cookie was used each week in part because the blinding may not have been completely effective. However, with a single subject being used in this study, it is difficult to extrapolate the significance of this study's results. [*This may have been one of the earliest suggestions by investigators that adverse effects of artificial food colors may possibly involve exaggeration of behaviors, such as irritability, which are not typically associated with childhood hyperactivity. Over subsequent years, this concept evolved into one of the shifts in some clinical studies from a focus targeted on artificial food colors as possible triggers of criteria-specific hyperactive symptomatology to a wider focus on artificial food colors more as provoking general behavioral disturbance (Bateman et al, 2004; McCann et al, 2007; Schab and Trinh, 2004; Weiss et al, 1980).]

(Rose, 1978)

Rose TL. The Functional Relationship Between Artificial Food Colors and Hyperactivity. J Appl Beh Anal, 1978 (Winter), 115 (No.4): 439-446.

Rose (1978) conducted a placebo controlled double-blind challenge study of a single artificial food color, tartrazine (Yellow 5), using a within-subjects design with a sample of hyperactive children (n=2) who were anecdotally reported by parents to be behaviorally responsive to the Feingold/KP elimination diet. Two diagnosed hyperactive children (female; 8 years) diagnosed with hyperactivity and maintained on the F/KP diet for at least 11 months as the apparently exclusive means of behavioral management were recruited for this study. The challenge vehicle was an oatmeal cookie with/without 1.2 mg tartrazine (FDC Yellow 5) (i.e., 0.05 mg/kg bw). The cookies were tested by 2 panels of adults and submitted to other testing verifying that placebo and color challenge cookies could not be differentiated. Throughout the study observers, children and parents were blind to treatment, but the coordinating experimenter was not blinded and neither parents nor children were specifically asked to verify whether they could tell the difference between color and placebo cookies. Procedurally, with the F/KP elimination diet being maintained throughout, the study was conducted over a period of 6 weeks with the following weekly treatment sequence: baseline, placebo treatment, color challenge, placebo treatment, color challenge, placebo treatment. After the 1-week baseline, subjects received one

cookie/day with breakfast throughout the remainder of the study. The active color challenge consisted of a single tartrazine cookie given on only one day during each of the two color challenge weeks. The sole behavioral outcome measure was the daily assessment by trained observers at school of specific behaviors characteristic of the hyperactive syndrome, including on-task duration, out-of-seat duration/frequency, and physical aggression frequency; these behaviors were monitored at approximately 3 hours after cookie ingestion.
Study Results: Visual inspection of the graphed data for each subject showed apparent exaggeration of hyperactive behaviors (increased out-of-seat behavior and decreased on-task behavior) associated with ingestion of the tartrazine containing cookie for both children. A randomization test showed these behavioral changes to be significant treatment related (tartrazine) effects. Treatment related changes in the frequencies of physical aggression were far less conclusive. Dietary infractions (a chocolate bar, tomatoes, a commercial poultry food item) produced similar behavioral changes. When dietary infractions did occur, the parents' observations of their children's behavior were reported to correlate perfectly with the behavioral data obtained by the independent observer. However, there is no information as to whether the parents' observations correlated with the observers behavioral evaluations for the double-blind placebo controlled challenges.
Study Assessment and Evaluation: The data from this study do appear suggestive of a functional relationship between ingestion of tartrazine on the behavior of two hyperactive diet-responsive children based on behavioral evaluations by independent observers in a school environment. However, the absence of any other outcome measure (e.g., parent ratings) to verify the observers' findings and the uncertainty regarding verification of effectiveness of the blinded conditions for the parents and children limits confidence in the reliability of these study results and the use of only 2 children as the study sample makes it difficult to extrapolate the significance of these suggestive findings.

(Williams et al, 1978)

Williams JI, Cram DM, Tausig FT and Webster E. Relative Effects of Drugs and Diet on Hyperactive Behaviors: An Experimental Study. Pediatrics, 1978, 61: 811-817.

Williams et al (1978) designed a clinical challenge study to focus on the relative effects of a diet free of artificial food colors, flavors and major preservatives in comparison to stimulant medication in managing hyperactivity in children. The study was conducted as a double-blind challenge trial comparing the effects of stimulant medication versus a mixture of artificial food colors in a heterogeneous group of children with hyperactivity or problem behaviors maintained on a modified Feingold diet [*Feingold's *modified* diet focused more on colors, flavors and preservatives and less on salicylates as substances most likely to trigger the hyperactive response (Feingold, 1975)]. The 26 children (24 male, 2 female) recruited for the study were 6 to14 years old, clinically assessed as hyperactive* based on parent and teacher reports, and responsive to medication for hyperactivity. [* Data collected during the study indicated that 7 children were apparently not hyperactive. Consequently, the study sample was behaviorally heterogeneous consisting of 19 hyperactive and 7 non-hyperactive children but with problem behaviors responsive to stimulant medication.] Unlike many other challenge studies involving artificial food colors, responsiveness to the Feingold elimination diet was not a criterion for admittance into this study. Two active treatments and two appropriate placebo/vehicle control treatments were used. The active treatments were: (1) chocolate cookies each with a mixture of nine artificial colors at a level of 1/2 the average daily intake of artificial colors (challenge was two

cookies per day) and (2) stimulant medication at dosages used prior to the study (challenge was two pills per day). Specific efforts were made to maintain the double-blind conditions of the study, including use of appropriate placebos, random assignment to treatment, challenge items provided daily by investigators, and having all study personnel and raters blinded to all treatments. Procedurally, the study period covered 14 weeks starting with a 1-week baseline with children on regular diets. During the next 5 weeks, all families were started on the modified Feingold diet which was then maintained throughout the rest of the study period. The primary challenge phase of the study was conducted over the next 4 weeks. On each challenge day subjects received a pairing of cookie (either placebo or color) and medication (either drug or vehicle) treatments. There were 24 possible complete sets of such paired-treatments. Each child was randomly assigned to one of those sets and rotated through the four paired challenge treatments (e.g., color cookie + drug, placebo cookie + drug, color cookie + drug vehicle, placebo cookie + drug vehicle), one paired-treatment each week. Subjects received either two color or placebo cookies and either two drug or placebo pills daily on the first 4 days of each week, i.e. Monday, Tuesday, Wednesday, Thursday). Generally, cookies as well as pills were taken in the morning and afternoon. This primary challenge period was followed by 2 weeks of preliminary data analysis to identify the two best treatment conditions for each child and then finally by an additional 2 weeks during which the children were rechallenged using only the two best treatments. The outcome measures consisted of behavior ratings by parents and teachers on treatment days (Conners 11-item checklist: parents daily, teachers alternate days), at the end of each treatment week (Conners 40-item checklist: parents and teachers), and at beginning and end of study (Conners 96-item checklist: parents only). Least square analyses were conducted for pairwise comparisons of means for checklist ratings.

Study Results: Based on grouped data analyses of the parent and teacher behavior ratings, the overall color challenge effects on the children's behavior were mixed. While the teacher ratings did reveal a significant main treatment effect with the color challenge adversely affecting the children's behavior, there was no confirmation from the parent ratings which found no significant main behavioral effects of the color challenge. In contrast to the inconsistent color challenge effects, both the teacher and parent behavior ratings found clearly significant main effects for stimulant medication in improving behavior, even though 7 of the subjects had such low overall behavior rating scores that they probably were not hyperactive. None of the data were not analyzed separately with and without the non-hyperactive children included. Based on analysis of data for individual subjects and a *criterion challenge response* defined as at least a 33% decrease in rating scores on placebo cookies compared to color cookies, parent ratings identified 3 *criterion responders* and teacher ratings identified 5 *criterion responders*. But, these findings were inconclusive since there was no agreement between the teacher and parent lists of responding subjects. Behavioral ratings were not conducted during the initial baseline period, when the subjects were still on their regular diets. Since no comparative behavioral data was available for children on regular diet and on the modified Feingold diet, no assessment can be made of the extent to which the modified Feingold diet may have affected (improved) the hyperactive/problem behaviors of the test subjects. Also, the results from the final 'rechallenge' phase of testing in this study using the 'two best treatment' were not presented.

Study Assessment and Evaluation: One general comment is a clarification in that the investigators' repeated reference to "diet effects" in presenting and discussing the study results is inaccurate and confusing, since the design of this study did not specifically assess the effects of "diet" itself on behavior. No comparison of behavioral scores (parent or teacher rated behaviors) was made between the baseline behaviors, when subjects were on 'normal' diet, and the study behaviors, when all subjects were maintained on the modified Feingold diet. Rather, this study

tested only the behavioral effects of challenge treatment (stimulant medication and cookies with artificial colors) in children being maintained on the modified Feingold diet throughout the study period. In terms of analysis of the data for the whole study group, while the effects of stimulant medication in improving the behavior of the test subjects were clearly significant, the overall effects of artificial color challenge on behavior were inconsistent. Teacher behavior ratings detected overall significant color challenge effects with worsening of hyperactive/problem behaviors but parent ratings provided no confirmation of any treatment related effects. Also, the results from analysis of data for individual subjects appeared to be inconclusive in that teacher and parent ratings each identified several children (5 by teacher and 3 by parent) as color responders but there was no agreement between parents and teachers as to which children were the responders. The investigators appeared to consider these differences between parent and teacher findings as real rather than perceived based on their view that the teacher ratings were more sensitive than the parent ratings due in part because the children received their color challenges in the morning and at noon and any color effects on behavior would have been more apparent within several hours after challenge during school hours but less apparent by the time the children returned home. This suggestion of an abbreviated time frame for color effects on behavior are not consistent with anecdotal reports by Feingold and parents that children on the Feingold elimination diet (free of artificial colors and flavors) exhibit dramatic behavioral deterioration lasting hours to days after even minor dietary infractions. Several additional limitations to this study were identified by the investigators: (1) behavioral checklists were the only outcome measure used, although two different sources of ratings (parent and teacher) were obtained; (2) the treatment of only 4 days limits confidence in inferring long term treatment effects; and (3) marked variations in rating scores possibly related to the 7 non-hyperactive children. The latter issue of variability could have been addressed by analyzing the parent/teachers ratings for the color challenges with/without the data from the 7 non-hyperactive children. Overall, in view of the clear effect noted by both parents and teachers of stimulant medication in improving hyperactive/problem behaviors, this reviewer views the disparity between the parent and teacher findings, in terms of the main findings of color challenge effects and the identification of different color responsive children, as definitely problematic for interpretation of the color challenge findings and considers the study to be suggestive but inconclusive and not supportive of an effect of artificial colors on hyperactive and problem behaviors in children.

(Conners et al, 1980)

Conners CK, Goyette CH and Newman EB. Dose-Time Effect of Artificial Colors in Hyperactive Children. Journal of Learning Disabilities, 1980 (November), 13(No.9): 48-52.

Conners et al (1980) conducted this challenge study to determine whether hyperactive children who appear to react to artificial food colors could show a pharmacologic dose-time effect by using more sensitive laboratory tests and observations than used previously. Nine diagnosed hyperkinetic/attention deficit children (5 male, 4 female; 5-10.5 years), previously reported by teachers and/or parents under non-blind conditions to exhibit marked behavioral improvement on the Feingold elimination diet and to show notable behavioral sensitivity to artificial colors, were recruited for this study. The challenge vehicle was chocolate cookies each with/without 15 mg of a mix of artificial colors (challenge consisted of 2 cookies/day). Procedurally, the study was conducted under double-blind conditions in 2 sessions at 1-2 week intervals, one session using color challenge cookies and the other session placebo cookies. The order in which sessions were

run was randomized for each subject (4 subjects given a color challenge session first and placebo session second, and 5 subjects tested with placebo first and color second. Each session consisted of baseline measurements, followed by a one-time consumption of 2 cookies (total color challenge was 30 mg) and then retesting at 45, 90, 135 and 180 min after challenge. Outcome measures at baseline and each test interval included subjective behavior ratings by the experimenter and objective tests including actometer readings, activity chair measures, and performance of a paired-associate attention/learning task. The paired-associate learning task was selected on the presumption that it would be relatively practice-free from session to session.
Study Results: No significant adverse effects of color challenge (30 mg of mixed colors) on behavioral of diet responsive hyperkinetic/attention deficit children were found for any of the dependent outcome measures. The children's level of activity appeared to increase with ingestion of both color and placebo challenges. The investigators speculate that this normal course of activity could mask any specific effects of color challenge. The paired-associate task showed a marked practice effect across sessions, which appeared to confound detection of any possible color challenge effects.
Study Assessment and Evaluation: In an attempt to explain the negative findings in this study, the investigators suggest several factors that may have minimized or confounded the detection of any significant color challenge effects, including the low detection power of the test with small numbers of subjects, the use of an ineffectively low level of artificial colors (30 mg) which greatly underestimates the more realistic levels of colors consumed by children, the increased levels of activity with ingestion of both color and placebo cookies, and particularly the marked practice effect for the paired-associated task across sessions. The suggestion that these factors may explain why treatment effects were not detected seem inconsistent with the fact that all of the children in this study were selected specifically based on teacher and/or parent reports of marked behavioral changes with the Feingold diet and notable sensitivity to artificial colors, and not merely subtle laboratory changes. This selection criterion was intended to enhance the ability to detect a real treatment related effect if there was such an effect. Within this context, it seems clear that this study failed to demonstrate a significant pharmacological dose-time effect for the 30 mg mixture of artificial food colors in hyperkinetic children.

(Conners, 1980a)

Conners CK. Artificial colors in the diet and disruptive behavior. In R Knights and D Bakker (Eds.) *Treatment of Hyperactive and Learning Disordered Children: Current Research*. Baltimore, University Park Press, 1980: 113-120.

This double crossover challenge study of the behavioral effects of artificial colors was conducted to replicate a prior study by Goyette et al (1978, II) but with a larger sample size of 30 children (assumingly, all were diagnosed hyperkinetic). The challenge vehicle was a chocolate cookie with/without a 13 mg mix of all approved artificial food colors (challenge consisted of 2 cookies per day). Efforts to blind the study were inherent in the use of color and placebo cookies that were made to be indistinguishable but no apparent efforts were made to ensure the effectiveness of the blinding for parent and teacher raters. Procedurally, this study was conducted over a period of 10 weeks consisting of an initial 2-week baseline period (regular diet), a 3-week period during which all subjects were place openly (non-blind) on the F/KP elimination diet, and a final 4-week double crossover challenge period during which those children showing a 25% or greater reduction in hyperactive symptoms were maintained on the F/KP diet and randomly assigned to one of two sequences of weekly challenge conditions (color/placebo/color/placebo or

placebo/color/placebo/color). During each challenge week children received either 2 placebo cookies or 2 color cookies, daily. The outcome measures consisted of Conners behavior ratings conducted by the parents daily within 3 hours of ingestion of the challenge cookies at the dinner meal and by the teachers on school days with their ratings based on behaviors exhibited during the first school period of the day.

Study Results: There were no significant adverse behavioral effects of challenge with 26 mg mixed colors administered daily for two alternating 1-week periods in the group of 30 hyperkinetic diet responsive children based on either parent or teacher behavior ratings.

Study Assessment and Evaluation: This study, which was intended as a replication of Goyette et al (1978), detected no significant adverse effects in hyperactive children of daily challenge with 26 mg mixed artificial colors. Since these negative results provided no confirmation of the color challenge effects reported in Goyette et al (1978), the reviewer concurs with the conclusion that the effects reported by Goyette et al (1978) were most likely due to chance and not biologically relevant.

(Swanson and Kinsbourne, 1980)

Swanson JM and Kinsbourne M. Food dyes impair performance of hyperactive children on a laboratory learning test. Science, 1980 (March 28), 207: 1485-1487.

Swanson and Kinsbourne noted that a number of earlier studies found no behavioral effects of artificial color challenge, when up to 26 mg of food colors were used as the challenge. Since they suggested that the lack of response in those studies may have been due to an insufficient low level of colors, Swanson and Kinsbourne (1980) designed the present placebo-controlled challenge study to assess the effects of larger challenge doses of artificial food colors in a hospital setting. The investigators tested 40 children (36 boys, 4 girls; average age of 10 years) who were referred to the hospital with problem behaviors suggesting hyperactivity. For 20 of these children the diagnosis of hyperactivity was confirmed (average Conners score of 16.2) and stimulant medication produced a favorable response (*hyperactive set*). The other 20 children were considered not to be hyperactive (average Conners score 12.3, which is below the criterion of 15 for a diagnosis of hyperactivity) and reported to have adverse response to stimulant medication (*non-hyperactive set*). None of the children were previously identified as responsive to the Feingold diet. The challenge item was a capsule containing either 100 mg or 150 mg of a blend of 9 artificial food colors; placebo capsules contained sugar. Ten (10) children from each study group, i.e. hyperactive and non-hyperactive, were randomly assigned to either the 100 mg or the 150 mg color challenge group. All medications were stopped prior to the study. Procedurally, all children were maintained on the Feingold diet for the 3 days prior to testing and also the 2 days during testing. In the morning of each of the 2 test days the children received either one placebo or color capsule (po), with the placebo/color order counterbalanced across subjects. On each day of testing the primary outcome measure, a paired-associate learning test (PAT) was administered 0.5 hour prior to treatment and again at 0.5, 1.5 and 3.5 hours after treatment. An additional outcome measure, the Conners behavior rating, was completed two times on each test day (neither the timing of the rating relative to treatment nor the identity of the rater was specified).

Study Results: Challenge with both 100 and 150 mg of a mixture of artificial food colors significantly impaired paired-associate test performance (deceased attention) relative to placebo treatment but only in the hyperactive set of children. Performance was not affected by color challenge relative to placebo treatment in the non-hyperactive children. The color effect in the

hyperactive children was significant at all post-challenge test intervals (0.5, 1.5 and 3.5 hours post challenge), indicating a rapid onset effect which increased in magnitude over time. Also, there was no significant difference in effect between the 100 mg and 150 mg color challenge levels, indicating no dose-response effect at these levels of treatment. The apparent effect of color challenge on performance of the paired-associate test was not supported by the findings of the Conners behavior ratings which were described as showing no differences between the color challenge and placebo treatment. Since none of the Conners rating data were presented nor were the analyses of these data described, it is assumed that the Conners ratings showed no overall color challenge effects and no effects in either group of children. Also, without the availability of the Conners rating scores it is not known whether maintenance of the children on the Feingold diet was effective in moderating the problem behaviors of either the hyperactive or non-hyperactive group of children during the study.

Study Assessment and Evaluation: The singular significant finding of this study was that challenge with a mixture of artificial food colors (both 100 and 150 mg levels) resulted in a significant deterioration in performance of a paired-associate test in a group of hyperactive children when compared with their placebo control performance. This color effect was not seen in the group of non-hyperactive children. However, the reported color challenge effect in the hyperactive children may be a questionable treatment effect and may possibly be an artifact due to an inexplicably unusual placebo performance of the hyperactive children (the placebo performance of the hyperactive children improved across test intervals but the placebo performance of the non-hyperactive children deteriorated across intervals). Confidence in the findings of a color challenge effect in the pair-associate test may be further limited by an uncertainty about the effectiveness of the blind in this study. Blinding procedures appeared to be limited to the use of sugar in capsules as placebo, random assignment of subjects to challenge dose group, and randomized order of placebo/challenge treatments but it is unclear whether the study personnel were blinded to treatment. Actually, the use of simple sugar capsules as placebo may not have adequately blinded either the children or study personnel to the colored challenge capsules. Also, no procedure was described to verify the blinding. In addition to the paired-associate test, the Connors behavior rating was another outcome measure used in this study and the results from this outcome measure provided no supportive evidence of any behavioral effects of the color challenge. The investigators attempted to dismiss the lack of effect with the Conners ratings by suggesting that the objective laboratory paired-associate learning test is more sensitive than the subjective Connors behavior ratings for documenting adverse effects of color on behavior. This argument seems inconsistent with the fact that these same investigators considered the Conners score sensitive enough to identify the study populations of subjects in this study as *hyperactive* or *non-hyperactive* and that other investigators have found significant treatment effects with parent/teacher behavior ratings but not with objective testing (Harley et al, 1978; Egger et al, 1985; Carter et al, 1993; Batemann et al, 2004). The absence of any dye-related changes in the Conners ratings in this study suggests that whatever, if any, effect the dyes may have on the PAT performance in the hyperactive children does not include eliciting an overall exaggeration of the hyperactive or other problem behaviors in either group of children. Overall, the questionable effectiveness of the blinding in this study, the absence of supportive evidence of any color related changes based on the Connors ratings, and the inexplicable difference in placebo performances of the paired-associate test by hyperactive and non-hyperactive children lessen confidence in the reliability of the reported color related changes in paired-associate test performance in the hyperactive children. In view of these limitations and caveats the reviewer considers the findings of this study to be inconclusive and to provide no credible support for the proposed association between hyperactivity or other problem behaviors

and artificial food colors.

(Weiss et al, 1980)

Weiss B, Hicks William J, Margen S, Abrams B, Caan B, Citron LJ, Cox C, McKibben J, Ogan D and Schultz S. Behavioral Responses to Artificial Food Colors. Science, 207 (March), 1980: 1487-1489.

Weiss et al (1980) conducted a clinical challenge study to determine if behavioral sensitivity (adverse effects) to artificial food color additives could be demonstrated in a controlled trial. It is notable that diagnosed hyperkinetic children were not the target population for this study. This study used 22 non-hyperactive children with problem behaviors (15 male, 7 female; 2.5 – 7 years), whom parents reported as responsive to open (non-blinded) treatment with an elimination diet that excluded artificial colors and flavors. In the design of this study these investigators were among the first to include parental rating of selected behaviors targeted for each child as one of the primary outcome measures in an effort to maximize detection of behavioral responses to artificial colors that may be specific to individual subjects. The challenge consisted of a soft drink with a 35 mg blend of 7 artificial food colors plus cranberry coloring. Placebo soft drinks contained a combination of caramel/cranberry coloring but no artificial colors. The two drinks were determined to be indistinguishable by sight, smell, taste, or stain color. Throughout the study all children were maintained on a version of the modified Feingold diet that excluded artificial colors, flavors, select spices and extracts, the preservatives BHA and BHT, and for some children designated fruits and vegetables. Procedurally, the study was conducted as a double-blind challenge over a period of 11 weeks with each child serving as its own control. At a specified time on each of the 77 days, each child consumed a bottle of soft drink. On 8 days distributed randomly for each child among weeks 3 through 10 of the study period, each child received the color challenge drink. No parent or member of the study team knew which child was being challenged on any given day. Parental ratings/observations comprised the main outcome measures. Prior to the study the each set of parents selected a 10-item checklist of target behaviors that served as individualized response criteria for their child throughout the 11-week experimental period. The checklist consisted of 7 aversive behaviors associated with dietary infractions and 3 positive ("good") behaviors. On each day of the study, the parents completed the target behavior rating two times, the first within 3.5 hours of treatment and the second at a later time. Twenty-four hours after each treatment parents also made a global estimate of each target behavior, recorded dietary infractions, kept sleep data, rated the day as a whole, commented on any deviations from daily routine, counted during the day occurrence of any of the 7 aversive behaviors, and completed the 10-item Conners Parent/Teacher questionnaire. The investigators did not consider this a "group" experiment but 22 separate experiments. Accordingly, statistical analyses were conducted on each child's data separately; no group data analyses were conducted.

Study Results: Only 1 of the 22 test subjects, a 3 year old girl, appeared to show a dramatic adverse response to the color challenge (significant worsening of 5 of the 7 aversive behaviors, the Conners behavior ratings, and all global behavioral measures). During the study, the comments made by the mother of this child showed that she was apparently able to identify correctly five of the eight color challenge days, which the investigators attributed to the mother's ability to discriminate the response to color. Although various procedures were used to blind the study, there was no indication that the parents or children were asked to verify that they could not tell the difference between the placebo and color challenge drinks. The remaining 21

subjects showed no convincing sensitivity to the color challenge, including one 3-year old male reported to exhibit an inconsistent increase in 2 aversive behaviors primarily on three of the eight challenge days but no treatment related changes in any of the other parental measures including the Conners ratings.

Study Assessment and Evaluation: This study reported that one out of 22 test subjects exhibited a dramatic adverse behavioral response to challenge with a 35 mg mixture of artificial colors. A second subject with only limited and inconsistent behavioral changes showed no convincing sensitivity to the color challenge. The investigators asserted that these data strengthen the evidence that "modest doses of synthetic colors, and perhaps other agents excluded by elimination diets, can provoke disturbed behavior in children.". However, the reviewer considers this conclusion to be clearly overstated. Some perspective must be given to the conclusions to account for the fact that the vast majority of subjects in this study, 21 of the 22 non-hyperkinetic children with problem behaviors selected for their reported positive response to diets free of artificial colors and flavors, did not show any convincing evidence of intolerance or any other type of adverse response to the color challenge. In this context one responder suggests a rare occurrence of adverse effect. Also, although notable efforts were made to maximize detection of treatment related behavioral effects by using individualized targeted behavioral ratings, the conclusions should reflect the fact that confidence in the relevance of the study findings is limited by the fact that the detection of treatment effects was based exclusively on only one source of behavioral assessment, the parents' observations. Use of only parents for assessing behavioral change in the children and not including for example trained observers or teachers does not allow for independent confirmation of any treatment related effects in the home or other settings. In the absence of any other source of outcome measures, it is difficult to assess the reliability of the parental reports of adverse behavioral changes as true treatment effects. Even accepting the parental findings as suggestive of color challenge effects, the more accurate conclusion would be that these data may strengthen the possibility that a small subgroup of children with problem behaviors that reportedly improve on a diet that eliminates artificial colors and flavors may exhibit an intolerance to modest doses of artificial food colors responding with adverse behavioral changes. Assuming some credibility to the parents reported improvements in problem behaviors when the children were placed on an elimination diet (e.g. Feingold), it is notable that 21 out of 22 of these children showed no convincing evidence of a significant sensitivity to the color challenge. This might suggest that, while a very small sub-population of children with problem behaviors may be intolerant to food colors, there may be various foods or food ingredients, other than colors, that are of possibly more significance in provoking or aggravating problem behaviors. In fact, a part of the conclusion by Weiss et al (1980) does state that " …. colors, and perhaps other agents excluded by elimination diets, can provoke disturbed behavior in children".

(Mattes and Gittelman, 1981)

Mattes J and Gittelman R. Effects of Artificial Food Colorings In Children With Hyperactive Symptoms: A Critical Review and Results of a Controlled Study. Archives of General Psychiatry, 1981, 38:714-718.

Mattes and Gittelman (1981) conducted a double-blind placebo controlled crossover challenge study to assess the behavioral effects of artificial food colors in a group of children heterogeneous for hyperactivity. Several elements were incorporated into the experimental design to maximize the likelihood of demonstrating any behavioral effects of the food colors.

Specifically, (1) the study included only children maintained on the Feingold diet because, as reported by parents, their behavior improved markedly on the diet and deteriorated quickly and dramatically with exposure to artificial food colors, (2) efforts were made to exclude placebo responders, and (3) high dosages of colors were used. This study was conducted using a group of 11 children (6 male, 5 female; 4-13 years) who were heterogeneous for hyperactivity (6 ADHD/ADD and 5 other problem behaviors or history of hyperactivity) and FKP diet responsive. The challenge items were cookies with/without a 13 mg blend of food colors (children were given an increasing number of cookies/day to a maximum of 6/day, i.e. a maximum color challenge of 78 mg). Placebo and color cookies were made to be indistinguishable in taste and appearance (Nutrition Foundation). All study personnel involved in evaluation and care of children were blind to treatment. Also, teacher, parent, psychiatrist and child were asked to guess the type of cookie at the end of each trial week (although not specified, it is assumed this was an additional means of gauging behavioral status and not to verify effectiveness of blinding conditions). Throughout the study, all children were continued on the same Feingold elimination diets they used previously. Procedurally, the study was conducted over a period of 5 weeks starting with a one week baseline. The second study week was a non-blind trial period of daily ingestion of placebo cookies to identify any children reacting adversely to placebo and exclude them from further study (two of 13 children originally recruited for the study were thusly eliminated). The third and fifth weeks of study were the experimental weeks during which the children were randomly assigned to receive either placebo or color cookies daily for an entire week each. Interposed between these two experimental weeks was a one-week (fourth study week) washout period without cookies. Each experimental week consisted of one cookie on the first day (i.e., a color challenge of 13 mg) and an additional cookie each day to a maximum of six cookies on the last two days (i.e., a maximum color challenge of 78 mg/day). Cookies were given at breakfast, lunch and after school. Several subjective and objective outcome measures were used including parent and teacher behavior ratings (Conners rating scale; a special scale developed to assess hyperactive symptoms; global severity of illness); psychiatric evaluation including overall change and a Children's Diagnostic Scale; and a laboratory test of distractibility with a psychologist's rating of behavior during the testing. All observations and testing were conducted during baseline and at the end of placebo week and each experimental week. The distractibility test was given 1.5 hours after ingestion of cookies. To evaluate timing of any behavioral reactions to challenge, teachers and parents also completed a brief Conners rating 3 and 5 days after the initiation of each type of cookie. At end of each experimental week, each child also completed a Brief Conners Scale and rated overall change and global severity.

Study Results: Three of the children failed to ingest the maximum 6 cookies per day. One child simply refused and two children consumed only four cookies due to severe behavior reactions one involving color cookies and the other placebo cookies. Overall, none of the behavioral measures, i.e, ratings, evaluations or laboratory distractibility testing by parents, teachers, children, psychologists, or psychiatrists, showed any significant differences between placebo and artificial color treatment. The absence of treatment related effects was also reflected in the fact that none of the participants guessed beyond chance the type of cookie being used in either treatment week. Repeated analyses of the study data excluding the data for children not diagnosed as hyperactive did not alter the negative results. Based on assessment of the data for individual subjects, the investigators reported that six children showed some difference between placebo and color cookie treatment but that the parents were the only raters who reported noticing this difference. All others (teachers, psychiatrists, psychologists) noted minimal if any differences, providing no confirmation of the responses noted by the parents.

Study Assessment and Evaluation: This double-blind placebo controlled challenge study using a variety of subjective and objective behavior outcome measures was designed specifically to maximize the likelihood of detecting significant adverse behavioral effects of artificial food colors in a heterogeneous group of children with ADHD/ADD or other behavioral problems, which reportedly improved on the Feingold elimination diet and dramatically deteriorated with exposure to artificial food colors. The results of this study showed that no significant adverse behavioral effects were produced by daily challenge to a blend of artificial colors for one week at dosages up to 78 mg per day. The reviewer concurs with investigators' conclusion that the results of this study indicate that artificial food colors do not adversely affect the behavior of school-age children with hyperactivity or other behavioral problems who are claimed by parents to be markedly sensitive to these agents. As noted in other reviews in this report, it may be assumed that there is some level of credibility to the parents reported improvement in problem behaviors when the children were placed on an elimination diet (e.g. Feingold) and dramatic behavioral deterioration with exposure to foods with artificial food colors. This suggested diet effect together with the fact that none of the subjects showed adverse response specifically to artificial colors under controlled challenge conditions might suggest that some foods or food ingredients, other than artificial colors, may be more significantly associated with provoking or aggravating problem behaviors in some children.

(Adams, 1981)

Adams W. Lack of Behavioral Effects From Feingold Diet Violations. Perceptual and Motor Skills, 1981, 52: 307-313.

Adams (1981) conducted a clinical double-blind, placebo-controlled challenge trial to investigate whether an *infraction* of the Feingold diet exaggerates behaviors often associated with hyperactivity in children. To minimize equivocal findings, this trial was designed to: (1) select hyperactive subjects who were reported to be behaviorally responsive to the Feingold diet and sensitive to violations of the diet and (2) evaluate treatment effects using primarily objective measures of activity, memory, attention and motor skills, in addition to parental ratings of behavioral change. A total of 18 hyperactive children were recruited for this trial (15 male/3 female, mean age of 7.6 years ranging from 4.4 to 11.6 years). The parents agreed that a positive diet effect was detectable in their child for at least three months and subjectively reported to have witnessed a prior negative effect following a diet violation using substances the same as or similar to those used as "challenges" in this controlled trial. The challenge item ("artificial snack") consisted of a chocolate cupcake with icing and a glass of lemonade both made from commercial mixes to which additional artificial red and yellow colors and artificial vanilla flavor were added. While the investigator stated that each serving of the "artificial snack" contained a total of 26.3 mg artificial food colors, only red (# 3; #40) and yellow (#3;#5;#7) colors were identified as being included. (There was a discrepancy in the publication's Table 2 between the stated total content per serving as being 26.3 mg and the total amount of itemized colors used which was 18.8 mg.) The amount of added artificial vanilla flavor was not specified and the other ingredients in the commercial mixes were not identified. The placebo ("natural snack") iced chocolate cupcake and lemonade were made with color/flavor-free natural ingredients consistent with the Feingold diet. While the parents were aware that one snack was "artificial" and the other "natural", neither the parents nor the observer knew which snack was given. However, the effectiveness of this blinding of the parents was not verified. Throughout the study, all children were continued on the same Feingold elimination diet they used previously.

Procedurally, each child was tested on two separate days with the placebo "natural snack" being given on one day and the challenge "artificial snack" on the other. The interval between test days was not specified. The children were randomly assigned to the placebo/challenge (n=9) or the challenge/placebo treatment order (n=9). On the first test day each child was given the scheduled snack 3 to 4 hours prior to the actual testing session. At the conclusion of that test session, the child was given an appointment for the second test session along with the snack for that session. In each test session the outcome measures consisted primarily of a battery of objective tests assessing overall activity level (activity room with multiple toys and games), fine-motor skill (drawing, handwriting), gross-motor skill (hopping, ball bouncing, bean-bag throwing), short-term memory (numerical and visual), and language (picture vocabulary). On each test day the parents also provided a checklist rating of their child's observed behavioral changes during the 3 to 4 hour period after the snack was consumed (no specific parental questionnaire was described).Within-groups analyses of variance and Chi-square analyses were performed. Data for individual subjects were not analyzed.

Study Results: No significant group differences between placebo and the "artificial snack" challenge treatments were found with any of the battery of objective tests or with the parental rating of changes in behavior. Also, no significant interaction effects, including age-related changes, were noted. Although data for individual subjects were not analyzed statistically, the investigator did mention that only small or slight changes in behavior were noted for any given child.

Study Assessment and Evaluation: Hyperactive children were specifically selected for this study based on their parents' anecdotal reports of notable behavioral improvement on the Feingold diet (eliminating artificial food colors, flavors, and natural salicylates) and behavioral deterioration with violations of that diet. Despite the very select nature of this group of children, an infraction of the Feingold diet under controlled clinical conditions did not produce any overall significant deterioration in a variety of objectively measured behaviors often associated with hyperactivity and did not result in any negative changes in behavior detectable by the parents. The investigator considered and dismissed two possible reasons for the absence of significant group treatment effects in this study, i.e. age and amount of color in the "artificial snack" challenge,. The investigator appeared to consider it unlikely that more pronounced effects might have been noted in this study if younger children had been tested, since no age-related interactions were found in the analysis of the data. Even though the total amount of food color used in the artificial challenge (26.3 mg) was somewhat lower than the average daily intake of colors for children (as estimated at the time of this study), the parents in this study assured the investigator that their children would show adverse behavioral changes if only part of the "artificial snack" challenge were consumed. It also appears that this trial focused more on verifying the possible adverse effects of an *infraction* to the Feingold diet, rather than specifically the possible adverse effects of artificial colors. While the colors represented a principle component of the "artificial snack", the iced cupcake/lemonade challenge contained added artificial vanilla flavor and were made from commercial mixes which undoubtedly contained other unspecified ingredients that were also considered *infractions* of the Feingold diet. An additional factor that may affect the results of a controlled challenge trial is the blinding of the study participants, since ineffective blinding may introduce an unintentional bias in behavioral assessments. Although the effectiveness of the blinding for the parents and children was not verified in this study, there was no reason to believe that the blinding may have been compromised or otherwise adversely affected the credibility of this trial's results.

Overall, the primary finding from this trial was that no significant overall adverse behavioral

effects in a group of hyperactive children resulted from a single *infraction* of the Feingold diet, based on a battery of objective tests and subjective parental ratings. Since the only artificial ingredients identified in the food items (iced chocolate cupcake and lemonade) used as the dietary *infraction* included 26.3 mg mixed artificial red/yellow colors and an unknown amount of artificial vanilla flavor, it may be concluded only that this study's findings provide no evidence of an association between ingestion of these particular artificial ingredients and behaviors associated with hyperactivity in a select group of school-aged hyperactive children. While this study does not disprove the possibility that there may be some negative behavioral effects produced by some artificial food substances or other food components or ingredients in some susceptible children, it does lend support to the views that, if diet is related in some way to problem behaviors, this effect is (1) less reliably demonstrated than parents predict and may possibly call into question the reliability of a parent's belief of what substances negatively affect their children's behavior, (2) involves fewer children than originally postulated, and (3) has a less dramatic and pervasive effect on attentional and motor performance than originally believed.

(Spring et al, 1981)

Spring C, Vermeersch J, Blunden D and Sterling H. Case Studies of Effects of Artificial Food Colors on Hyperactivity. The Journal of Special Education, 1981, 15 (Number 3): 361-372.

Spring et al (1981) conducted a double-blind, placebo-controlled double-crossover clinical trial to test the Feingold hypothesis that artificial food colors cause hyperactivity in some children. To maximize the probability of demonstrating an unequivocal causal relationship between food colors and hyperactivity, the investigators recruited only children who were already on the Feingold diet and whose mothers reported a diet-related reduction of hyperactivity symptoms. Additionally, to ensure that any treatment related findings were causal and not coincidental effects, a replication procedure similar to the original experimental procedure was also carried out. Six male children (8 to 13 years), all considered hyperactive and several with learning problems and neurological signs, were included in the trial. Although no mention was made of a formal diagnosis of hyperactivity, mothers used a hyperactivity scale and rated their child's behavior before starting the Feingold diet at scores ranging from 23 to 27 and after starting the Feingold diet at scores ranging from 12 to 15. This rating of behavior before the diet was based on the mothers' recollections, since the approximate time before this study that these children were on the Feingold diet ranged from 7 to 32 months. The test materials were chocolate cookies (Nutrition Foundation). Placebo (control) cookies contained no artificial colors. Each active challenge cookie contained a 13 mg mix of all eight approved artificial food colors and, since subjects received two cookies each test day, the total daily challenge was 26 mg. A pilot study with college students confirmed that the placebo and challenge cookies could not be discriminated by sight or taste. The Feingold diet was continued throughout the 8-week primary trial and the subsequent 6-week replication procedure. Procedurally, the initial 2 weeks of the primary trial were baseline with no cookies administered but hyperactivity ratings taken as scheduled (see below) for both weeks. The subsequent 6 weeks constituted the primary experimental period during which two cookies were given daily on Tuesdays, Wednesdays and Thursdays of each week, one cookie before school and one cookie after school. Active challenge cookies only were given during two of these experimental weeks and placebo cookies only during the other four experimental weeks. The children were randomly assigned to one of 2 groups. The three children in Group 1 received the active challenge cookies during experimental

weeks 1 and 4, and the three children in Group 2 were challenged during experimental weeks 3 and 6. On each day that a child ate cookies, both the mother and teacher were asked to rate their child's behavior using two outcome measures. One measure was an abbreviated 10-item modified version of the Conners hyperactivity rating scale, which the investigators verified with an internal assessment and concluded was a valid measure of hyperactive behavior. However, since this was a modified Conners scale, the score level that indicated hyperactivity was unclear. The second outcome measure was a global behavior judgment whereby the respondent (mother or teacher) was to guess whether the child had received the active challenge or the placebo based on the respondent's assessment of the child's general behavior on that day. Both ratings were collected by telephone each day but it is not clear whether reporting times for mothers and teachers were standardized to a defined period of time after ingestion of the before-school and after-school cookies. Diet diaries were also collected on test days to assess nutrient intakes on the Feingold diet but differences appeared to reflect individual food preferences. Each of the 6 subjects constituted a single subject experiment. The hyperactivity ratings data were analyzed by computing point-biserial correlations separately for mothers' and teachers' ratings. The global behavior judgment (daily guesses) data were analyzed with Fisher exact-probability tests. Analysis of the grouped data was not conducted. Subsequent to the original trial, a replication procedure was carried out to confirm any treatment related findings. This replication procedure was similar to the original trial, except that baseline lasted only 1 week, cookies were given for 5 weeks with active challenge cookies being given during treatment weeks 2 and 4, and only the mother's rating data (hyperactivity rating and global behavior judgment) were collected.
Study Results: When the data were analyzed with both active challenge periods combined, the global behavior judgment measure in which mothers and teachers guessed which treatment their children received revealed that significantly accurate guesses were made only by one subject's (Subject E) teacher. The guesses of Subject E's mother, as well as the guesses of the remaining subjects' mothers and teachers were not significantly accurate. Analyses of the daily hyperactivity ratings across both challenge periods showed Subject E's mother and teacher ratings of increased hyperactive behaviors to be significantly correlated with active challenge (although, as noted below in the *Study Assessment and Evaluation* section, when the rating data were viewed for each week, the ratings by both mother and teacher were found to be inconsistent between challenge periods and the raters did not agree on which challenge period the change in behavior occurred). Additionally, the overall hyperactivity ratings of Subject A's mother, but not his teacher, were also significantly correlated with color challenge. Data for the remaining subjects indicated little relationship between overall daily hyperactivity ratings and daily treatment.

For a more in depth assessment, the hyperactivity ratings data for Subjects A and E were viewed separately for each week of the study period (these data presented graphically in Figure 2 of Spring et al, 1981). Both Subject A and E were challenged during experimental week 3 (first challenge period) and week 6 (second challenge period). With weeks viewed separately the ratings of Subject A's mother inconsistently showed a peak increase in adverse behaviors during the first challenge period but not during the second challenge period. Furthermore, it was revealed that during the first challenge period a home event occurred that may have influenced the mother's rating and, consequently, the investigators considered it likely that the relationship between the peak increase in adverse behaviors of Subject A and the first color challenge may have been coincidental rather than causal. The weekly hyperactivity scores for Subject E showed, what the investigators viewed as, peaks in the teacher's ratings during both challenge periods, which coincided with a peak in this subject's mother's ratings but only during the

second challenge period. Based on their assessment of these data, the investigators considered it reasonable to (tentatively) regard the positive relationship between Subject E's diet challenges and his increase in adverse behavior as causal. (As will be discussed in the *Study Assessment and Evaluation* section below, this reviewer considers that this interpretation of the study data does not reflect an adequate assessment of the mean weekly hyperactivity ratings for Subject E.)

Acknowledging the possibility of a coincidental relationship between food colors and hyperactivity for Subject E, the investigators conducted a replication procedure with this subject. The replication procedure, based only on the mother's daily hyperactivity ratings and global behavior judgment (daily guesses about treatment), failed to confirm any significant effect of the active (color) challenge for Subject E.

Study Assessment and Evaluation:

In the analyses of the mothers' and teachers' daily hyperactivity ratings and global behavior judgments (guesses about treatment) with data for both active challenge periods combined, only 2 (Subjects A and E) of the 6 reportedly diet-sensitive subjects initially appeared to show any indication of an adverse treatment related effect. For Subject A, with both active challenge weeks combined the singular significant indication of a possible treatment effect was in the mother's increased hyperactivity ratings. However, when these data were viewed for each week individually, the mother's hyperactivity ratings were actually shown to be inconsistently increased only during the first active challenge period but not during the second active challenge period. Furthermore, due to an undisclosed home event the apparent association between the behavioral changes and color challenge reported for Subject A was considered coincidental and not causal, thereby dismissing the presence of any treatment related effects for Subject A. For Subject E, with both active challenge weeks combined both the mother and teacher reported significantly increased adverse behaviors in the daily hyperactivity ratings but only the teacher's global behavior judgment (guess) about treatment was significantly accurate. When the hyperactivity ratings data for Subject E was viewed separately for each week of the study period, the ratings by both mother and teacher were found to be inconsistent between challenge periods and the raters did not agree on which challenge period the change in behavior occurred. The mother's ratings showed a peak in adverse behaviors during the second challenge period but no changes in the ratings during the first challenge period. Contrarily, Subject E's teacher's ratings showed a peak in adverse behaviors during the first challenge period but, in the opinion of this reviewer, showed no adverse changes in rated behavior during the second challenge period. The latter opinion, which conflicts with the investigators' interpretation, is based on the fact that the teacher's hyperactivity rating score (approximately 8) for Subject E during the second challenge period was lower than the placebo week 2 hyperactivity score (approximately 10). Consequently, this reviewer does not consider the teacher rating during the second challenge period to indicate a treatment related change. The absence of any treatment effects on the behavior of Subject E was confirmed in the replication procedure. In assessing the overall findings of this study it should also be noted that a reasonable effort was made to confirm the blinding of the challenge and, although the effectiveness of blinding was not specifically verified for parents, teachers or children, there was no indication that the participants held negative or biased attitudes about the Feingold hypothesis or that the credibility of the outcome measures was in any way affected.

Overall, in deference to the investigators' conclusion that this study (primarily based on the primary trial data for Subject E) provides equivocal support for Feingold's hypothesis, this reviewer's assessment concludes that this study's findings provide no consistent or credible evidence that hyperactive behavioral effects were elicited or exaggerated by repeated exposure to

26 mg mixed artificial food colors during two separate weeks in a small sample of hyperactive boys. It is possible that these children, selected for their reported sensitivity to the Feingold diet, could have been responsive to food elements in that diet other than artificial colors.

(trials conducted from 1982 to 2007)

(Salamy et al, 1982)

Salamy J, Shucard D, Alexander H, Peterson D, Braud L. Physiological changes in hyperactive children following the ingestion of food additives. Int J Neurosci. 1982 (May), 16(3-4):241-6.

Salamy et al (1982) conducted a study to explore a method for assessing dietary interventions based on each individual's response irrespective of group performance. Their approach used a single case experimental design based on physiological measures of autonomic nervous system (heart rate or HR) and central nervous system (electroencephalogram or EEG) activity for evaluating the effects of food additive infractions of the Feingold elimination diet in hyperactive children. The study was conducted on a total of 8 children (5-11 years). Four of the children were previously diagnosed hyperactive and reported to be responsive to the Feingold diet. The other 4 subjects were non-hyperactive sibling controls. All children were currently being maintained on the Feingold diet and were continued on this diet throughout this study. The challenge item containing artificial additives was a lemon flavored commercial drink colored with Red 40 (amounts of additives including Red 40 were not specified). The placebo was a home-made lemonade drink colored with beet juice. Procedurally, the double-blind placebo controlled randomized crossover study was conducted on two separate test days one week apart. On each test day baseline physiological recordings (HR, EEG) were taken for 10 minutes. Immediately after baseline, children were given their pre-assigned placebo or color drinks orally through a straw from an opaque container and served along with popcorn. Neither the investigators nor the children knew which drink was given on either test day. Following administration of the drink, children played for one hour, after which physiological measurements were again recorded for 10 minutes. (The parents of two hyperactive children reported marked behavioral changes after the children left the laboratory. Even though it was not stated which treatment (placebo or color) these children had just received, the investigators gave these children an additional day of testing to determine whether a post treatment period longer than one hour may be necessary to observe treatment effects. In this retest physiological recordings were obtained four hours after a double dose of the color drink, but the double-blind design was abandoned for this additional test, confounding the reliability of these retest results.)
Study Results: Based on measures of HR and EEG, there was no evidence of significant physiologic changes (EEG or HR) specifically attributable to the color/artificial additives treatment in either hyperactive or control subjects. The most striking finding was a relative absence of physiological changes seen in the control subjects but an abundance of physiologic changes in the hyperactive children under both placebo and color/additive treatment conditions. The hyperactive children appeared to be influenced by the placebo as readily as they were by the drink containing color/additives. Thus there appeared to be a difference between groups (hyperactive vs sibling control) but not between treatment conditions (color/additives vs placebo). These findings suggest that the hyperactive children compared with controls were simply more labile physiologically, responding generally to the experimental situation.
Study Assessment and Evaluation: The data showed no reliable evidence of significant physiologic changes (EEG or HR) specifically attributable to the color/artificial additives

treatment in any of the 4 hyperactive or 4 control children. However, while suggestive, this negative finding should be considered inconclusive. One of the more apparent reservations about the reliability of this negative finding is that no information was given in this study about the levels of Red 40 or other artificial additives present in the active treatment. The possibility that the levels of color/additives used may have been too low to elicit any physiologic response cannot be dismissed. Additionally, confidence in these results is also limited by the fact that only one recording session was made with each experimental treatment, i.e. placebo and color/artificial additives, for each test subject. Variability in physiologic responses between and within subject to any treatment is expected. Confidence in the detection of real treatment related physiologic effects or in the determination of no treatment related effects relies on repeated testing of the experimental treatment to replicate the physiological response or verify the absence of a treatment related effect. Consequently, the inconclusive findings from this study do not provide sufficient evidence to conclude that there were no treatment related physiologic effects of challenge with Red 40/artificial additives in hyperactive or control children.

(Thorley, 1984)

Thorley G. Pilot study to assess behavioural and cognitive effects of artificial food colours in a group of retarded children. Dev Med Child Neurol, 1984 (February), 26(1):56-61.

Based in part on Thorley's views (1980) that some clinical studies do suggest that subgroups of hyperactive children may benefit most from additive-free diets (Goyette 1978; Williams et al ,1978; Swanson and Kinsbourne, 1980), that adverse effects of artificial colors are more apparent at higher, than lower, doses (Swanson and Kinsbourne, 1980) and that only small numbers of "normal" children, i.e. w/o psychiatric disorder, show adverse behavioral effects after color challenge (Rose 1978, Weiss 1980), Thorley (1980) conducted a clinical challenge study designed to maximize detection of behavioral and cognitive response to color challenge. The following features were incorporated into this study: (1) use of high challenge dose levels, i.e. 2 times the estimated daily intake; (2) use of an optimum post-challenge time for testing; (3) use of a 2-week washout period and no cross-over challenges to minimize the confound of long-lasting behavioral effects of colors; (4) use of sensitive measures; (5) use of a residential setting to allow control of study conditions; and (6) use of a relatively homogenous population with all children intellectually retarded. The study was conducted in a residential school using 10 intellectually retarded children (8 male, 2 female;12 years average age) exhibiting inattentive/hyperactive behavior. The children were not previously identified as diet responders. The challenge material was a cocoa drink with/without a 91.8 mg blend of 16 approved artificial food colors (UK) (twice the estimated daily intake). Although a placebo control was used and all participants were blinded to challenge (children, teachers, care staff, and psychologist), there was no determination made of the effectiveness of the blinding procedures. Procedurally, the experimental design was a placebo controlled double blind challenge study (without cross-over) conducted over a period of 28 days with children maintained on an elimination diet (modified Feingold diet) free of all colors, flavors and preservatives throughout the study. The initial 14 days of the study was a "washout period" during which the elimination diet was started and children were 'habituated' to the placebo cocoa drink (assumed this was given daily during he "washout period"). This was followed by a 14 day "challenge period" in which 2 randomly selected children at a time received the color challenge drink for 2 consecutive days at 9:30 AM and all other children received the placebo cocoa drink. Outcome measures consisted of multiple

subjective and objective testing, including: (1) teacher behavior ratings in morning (Conners Teacher Scale and a Devised Individual Rating Scale of 5 most problematic behaviors); (2) Care Staff behavior rating (Conners Parent Scale) at meal-time and recreation times; and (3) Psychologist psychometric testing one hour after challenge (Porteus Mazes for visuomotor skills and Paired Associates learning test for short-term verbal memory) and actometer motor activity. The data for each measure were simply summed to arrive at group scores for the baseline condition (2 weeks prior to challenge) and the color challenge condition (2 days of color challenge). The two sets of data for each parameter were simply compared using uncorrected multiple t-testing.

Study Results: Although the study was designed to maximize detection of any behavioral or cognitive response to artificial color challenge, there were no significant treatment related effects found for any behavioral or psychometric test measure in a group of intellectually retarded children at a high challenge dosage of 91.8 mg of artificial food colors, which is twice the estimated daily intake. There was no information presented to determine whether the elimination diet, relative to the regular diet, may have affected the disruptive behavior of the children.

Study Assessment and Evaluation: Since no convincing adverse effect of artificial colors was found in a study designed to maximize detection of behavioral and cognitive responses to challenge with artificial food colors, the Reviewer agrees with the investigator's conclusion that this study seems to indicate that it would be unlikely that retarded children as a whole would show gross adverse effects to artificial food colors. Even though there may have been some question regarding the effectiveness of blinding and the appropriateness of analyzing the data using uncorrected repeating t-testing, there appeared to be no apparent effect of these factors on the study results.

(Egger et al, 1985)

Egger J, Carter CM, Graham PJ, Gumley D and Soothill JF. Controlled trial of oligoantigenic treatment in the hyperkinetic syndrome. Lancet. 1985 (Mar 9), 1(8428):540-5

Based on the possibility that food allergy may play a role in causing or aggravating the hyperkinetic syndrome in children, Egger et al (1985) conducted one of the first multistage studies to treat overactive (hyperactive) children with a diet that contains few varieties of foods (oligoantigenic diet), to identify in open trial possible provoking foods in those children with improved behavior, and then to test the effects of those presumptive provoking foods in a randomized, placebo-controlled, crossover challenge trial. The subjects in this study were atypical hyperactive children, either diagnosed hyperkinetic or exhibiting problem behaviors including overactivity, and many with neurological disorders, allergic conditions, or other associated physical conditions. All subjects were in a special clinic to assess and treat overactive children by dietary means. Of the 76 children placed on an open (non-blinded) oligoantigenic diet, 62 (82%) were found to be "food intolerant" with noted improved behaviors on the diet. Thirty-one of these "food intolerant" children (23 males, 8 females; average age 7 years) were included in open re-introduction trials (one test food per week) to identify provoking foods and 28 completed the double-blind placebo controlled challenge trials with each child being challenged with one of their provoking foods. In the open re-introduction all children reacted to more than one food but the most common food ingredients causing reactions were benzoic acid and tartrazine but no child was sensitive to these food additives alone. The varieties of food used as challenge items in the double-blind trials included cow milk, orange juice, wheat, tinned food,

tartrazine (150mg) in capsule, or benzoic acid (150mg) in capsule (8 subjects received either tartrazine or benzoic acid). Placebo items, respectively, included soya, sheep milk, or goat milk; blackcurrant, apricot or pineapple juice; rice or oats; and ascorbic or calcium tartrate. Staff members verified no difference in taste between active and placebo treatments and all study participants were blind to the order of challenge, but no verification of blinding effectiveness for the families was carried out. Procedurally, the double-blind challenge phase of this study was conducted over a period of 4 – 6 weeks during which children were maintained on their individual accepted oligoantigenic diet. Each child was given either his/her incriminated food item or placebo for 1 to 2 weeks and an additional 1 to2 weeks with the opposite treatment; there were at least 2 weeks of washout between treatments. The order in which each child received the incriminated food or placebo was randomly determined. The outcome measures included: (1) parent behavior ratings completed daily (Conners abbreviated scale); (2) parents and clinician made overall behavior assessments at end of each treatment period; and (3) psychologist conducted overall behavior assessments and psychological testing at end of each treatment period (included actometer measures, matching familiar figures, Porteus maze, and intelligence assessments – note that not all tests were done by all children). (Skin prick testing and serum Ig E level determinations were made during the initial open testing of the oligoantigenic diet and not during the challenge phase of testing).

Study Results: In double blind challenge, 'overall behavior assessments' by parents, psychologist and clinician (all $p<.001$) linked better behavior with the placebo challenge significantly more often than with the active food challenge. No significant order effects occurred. Parent behavior ratings using the Conners scale showed similar trends with active provoking food challenge eliciting more hyperactive behavior than placebo ($p<.01$). However, while there were no significant order effects for the parent ratings, data analyses were complicated by a significant treatment order interaction involving the placebo values for the two treatment orders. In contrast to the behavioral ratings/assessments, none of the objective psychological tests (psychologist ratings, actometer measures, match familiar figures, and Porteus maze) showed any significant differences between active provoking food challenge and placebo, although non-significant trends did generally indicate placebo behavior better than active challenge. Atopy status did not differ between the diet responders and non-diet responders.

Study Assessment and Evaluation: : In open trials of the oligoantigenic (few foods) diet 62 out of 76 (82%) overactive/atypical hyperkinetic children appeared to show improved behaviors, suggesting possible food intolerance. Atopy (allergic sensitivity) was not different between the diet responders and non-responders, but IgE levels were higher for the diet responders. Multiple provoking food items were identified in open trials. All children appeared to react to more than one food and the most common food ingredients causing reactions were benzoic acid and tartrazine, but no child was sensitive to these food items alone. Select incriminated food items were used to test 28 responders under double-blind challenge conditions. Overall assessments by multiple sources including parents, psychologist and clinician, linked various provocative food items with increased hyperkinetic/overactive behaviors in children and the parents' daily behavior ratings also indicated the same association. In the presentation of results graphic representation of some of the behavior data indicates that several children (approximately 6) did not respond to the provoking food challenge. However, the investigators did not specify what provoking food items these ‘non responding’ children received. So it is unclear whether any of these “non-responders” were among the 8 children challenged double-blind with either tartrazine or benzoic acid. Also, there appeared to be an inexplicable but not significant difference in placebo values between the treatment order P/A (lower placebo values) and the treatment order

A/P (higher placebo values). However, this appeared to have no effect on the challenge related behavioral effects. In contrast to the subjective measures, a battery of objective psychological tests, including an actometer measure of motor activity, showed no confirmation of any provoking food challenge effect on hyperactive behaviors. In contrast to the subjective rating and assessment measures, none of the objective psychological tests showed any significant adverse behavioral effects with the suspect provoking food challenge, although non-significant trends did generally indicate placebo behavior better than active challenge. Since significant adverse effects of provoking food items was corroborated by several sources of outcome measures (parents, clinician and psychologist), it seems likely that the failure to detect treatment related effects with the objective psychological test in this study may have been due to the possibility that those tests were not sensitive enough to detect either the type or magnitude of behavior changes specifically associated with food intolerance. Although there is some uncertainty regarding validation of the effectiveness of the blind for the parents and children, this study does provide sufficiently reliable findings to indicate that some children with problem behaviors may be intolerant to a variety of food items, including but certainly not limited to artificial food colors and preservatives.

(David, 1987)

David TJ. Reactions to dietary tartrazine. Archives of Disease in Childhood, 1987, 62: 119-122.

David (1987) conducted a double-blind challenge study of selected children in a hospital setting to verify observations by parents of a definite history of an immediate adverse behavioral reaction (within two hours) to a food additive, particularly tartrazine and benzoic acid. This study used 24 children (19 male, 5 female; 2-12 years) with problem behaviors related to hyperactivity who were maintained on elimination diets that avoided food additives and any other suspect foods. Although the extent to which the elimination diets diminished problem behaviors appeared to be variable among the children, all parents reported the children to exhibit obvious adverse behavioral reaction within two hours of any lapse of the elimination diet. For all subjects parents reported tartrazine to be the additive that caused the most severe adverse reaction; benzoic acid was reported to cause similar adverse effects but in fewer subjects. Twelve of these subjects were inpatients in the general pediatric ward and 12 came to the ward as outpatients. The challenge items consisted of either orange juice or blackcurrant (Ribena) drinks with/without 50 mg and 250 mg tartrazine (Yellow 5) and separate drinks with/without 50 mg and 250 mg benzoic acid. All drinks (in this study) were taken through a straw because of the characteristic staining of the skin around the mouth after drinking solutions containing tartrazine. The investigator noted the difficulty of disguising tartrazine, particularly at high concentrations which give a characteristic appearance to food and drink that is quite obvious even in highly colored food or drink. The investigator considered capsules as not suitable because of the obvious bright color of those containing tartrazine. Parents of the outpatients knew on which days challenges were to be performed, but for the inpatients neither the parents nor the nursing staff knew on which days the challenges were to be performed. Procedurally, all children were maintained on their special diets throughout the study. Each child was given the placebo vehicle drink daily throughout the period of admission. The duration of this placebo treatment was not specified. On the first experimental day a member of the medical staff gave each child a challenge drink with 50 mg tartrazine and behavior was monitored for at least two hours. A second challenge drink with 250 mg tartrazine was given at least two hours after the first

challenge and again behavior was monitored for at least two hours after treatment. It is not clear whether placebo treatments were also given on challenge days. Subsequently, but on a separate day, all subjects were challenged with the two dosages of benzoic acid drinks following the same procedure as was used for the tartrazine challenges. It is not clear how many days were allowed between tartrazine and benzoic acid treatments and whether placebo treatments were given on those intervening days. Apparently, no randomized treatment order was used, either between tartrazine and benzoic acid challenge periods or between each of their dosages. The outcome measures consisted of each child being observed after placebo treatment and active challenges by a parent and by the nursing staff, each of whom were asked to keep a timed record of any change in the child's behavior. No specific or structured scoring system was used and the duration of observation was not specified. No procedures for statistical analysis of the data were described.
Study Results: No child (inpatient or outpatient) was reported to show any change in behavior as noted either by the parents or the nursing staff after administration of placebo or a single oral treatment with 50 mg or 250 mg of either tartrazine or benzoic acid. However, only a summary description of these findings were presented. No itemized data were presented and no data analyses reported.
Study Assessment and Evaluation: This double blind challenge study using rather high dosages of 50 and 250 mg tartrazine and benzoic acid did not appear to produce any observable adverse behavioral reactions in children who were reported to have a definite history of immediate adverse behavioral reactions to these food additives. However, several prominent limitations and weaknesses in the design and conduct of this study should be noted including the facts that the measure of behavioral change was based solely on unstructured observations by parents and nursing staff, that randomized treatment conditions were not used, that half of the parents were not blind to active challenge days, and that itemized observation data were neither presented nor statistically analyzed. These factors raise questions about the sensitivity and reliability of this study in detecting more than overt treatment related effects. Within the context of these caveats this study, even at high challenge dosages of tartrazine and benzoic acid, provides no support for the parents' anecdotal reports of immediate and obvious adverse behavioral reactions specifically attributable to these food additives. Since all of the children in this study were on food elimination diets that avoided a variety of food items, it may be possible that the previous parental reports of obvious adverse behavioral responses in these children may have been associated with an intolerance to food items other than tartrazine and benzoic acid. However, this study cannot exclude the possibility that there may be some subgroup of children who have more specific or subtle behavioral effects from ingestion of additives such as tartrazine and benzoic acid.

(Rowe 1988)

Rowe KS. Synthetic food colourings and 'hyperactivity': a double-blind crossover study. Aust Paediatr J, 1988 (April), 24(2):143-7.

Rowe (1988) conducted a two-phase clinical study designed (Phase I) to identify a group of children who were suspected hyperactive and reportedly improved in an open trial of the Feingold diet and (Phase II) to assess the relationship between certain behavioral characteristics and ingestion of two commonly implicated artificial food colors, tartrazine (Yellow 5) and red carmoisine. In Phase I, 55 heterogeneous children, both with and without suspected "hyperactivity", were tested in an open (non-blinded) trial of the Feingold diet. While 40 of those children (72%) appeared to show behavioral improvement on the diet (26 due to a possible placebo effect), only 14 (25%) were suspected of reacting adversely to the ingestion of foods

containing additives, particularly the red and yellow artificial food colors. From this latter group of suspected reactors 9 children (7 males, 2 females; 3-15 years) were selected for Phase II double-blind testing. The challenge items were colorless capsules with or without 50 mg of either tartrazine (Yellow 5) or red carmoisine in an inner (capsule) surrounded by lactose. The placebo inner capsule contained lactose. All children were maintained on their Phase I additive-free diets throughout the study. Procedurally, one capsule was given to each subject daily at breakfast for the entire study period of 126 days (18 weeks). After variable placebo lead-in periods of 3, 4 or 5 weeks, tartrazine and carmoisine capsules were each given daily for one week on 2 separate occasions (i.e., a total exposure of 2 weeks for each color) with placebo washout periods of 2 or 3 weeks between each week of color challenge. The sequence of color challenge varied for each subject (assumed randomly). Parents, subjects and the data collector were blind to this challenge sequence. The behavioral outcome measure consisted only of a daily behavior rating completed by the parents, using a "frequent symptoms" 8-item checklist. (Attempts to obtain teacher behavior ratings were incomplete and did not provide a source of corroboration for any treatment effects.) The "frequent symptoms" checklist was compiled from behaviors commonly reported by parents as being associated with exposure to foods containing artificial colors. Data for each child were analyzed separately.

Study Results: One of the 9 children recruited for this study did not complete the challenge phase of testing. Two of the remaining 8 children who were heterogeneous for hyperactivity, claimed to be responsive to the Feingold diet, and suspected of reacting adversely to foods with synthetic additives (commonly red/yellow colors) were shown by parental ratings ('frequent symptom' behaviors) to show significant adverse behavioral responses to double-blind challenge with either 50 mg tartrazine or 50 mg carmoisine. One color responder was a 7-year old non-hyperactive female and the other was a 12-year old hyperactive male. Both responded with increased activity, irritability, low frustration tolerance, sleep disturbance and short attention span. Their responses began within hours after initiating each color challenge treatment and continued until 3-4 days (female) and 3.5 weeks (male) after the last of each color challenge treatment. Both children were atopic, but their symptoms of asthma and allergic rhinitis were not exacerbated by the food colors. The data for the remaining 6 children yielded results which were neither consistent nor statistically significant.

Study Assessment and Evaluation: Two (one hyperactive, one non-hyperactive) out of 8 children heterogeneous for hyperactivity with problem behaviors, responsive to the open trial with the Feingold diet and claimed to be behaviorally reactive to foods with artificial additives (commonly red/yellow colors) were shown by parental ratings ('frequent symptom' behaviors) to respond adversely to controlled double-blind challenge with either 50 mg tartrazine or 50 mg carmoisine. Confidence in the reliability of these positive findings is limited by the absence of any other source of behavioral measures to corroborate the parental ratings (teacher ratings were incomplete) and by the lack of attempt to verify the effectiveness of the blind for parents and children. At best, the study's finding of 2 possible color responders may suggest that a poorly characterized small subpopulation of children with behavioral problems may be intolerant to certain artificial food colors. Since the six children who did not respond to the controlled color challenges were reported in Phase I to have improved on the Feingold elimination diet and to have reacted to foods containing artificial colors, it may be possible that these Phase I behavioral responses may have been associated with some intolerance to food items/components eliminated from the Feingold diet other than artificial colors. Interestingly, the investigator points out that only one of these two possible color responders originally showed any signs of "attention deficit" behavior and that the common features of behavior associated with ingestion of food colors, as described by parents, emphasized extreme "irritability", "restlessness" and "sleep disturbance"

rather than those associated with "attention deficits" (implying rating scales designed to focus on hyperactivity related behaviors may not be sensitive to the range of other behaviors more typically exhibited by color responders). The author viewed this as suggesting that the behavioral dimensions associated with the ingestion of synthetic food colors require careful identification, since the inclusion of children in clinical trials on the basis of 'attention deficit' alone may miss some color reactors. This viewpoint effectively broadens the suggested behavioral impact of artificial food colors from only children with ADHD and related behaviors to children with more general types of problem behaviors, including but not limited to those associated with ADHD. Within this context, while this study does not support a causal association between artificial colors and 'hyperkinetic disorder' (ADHD) in children, the study results in fact suggest that the small subpopulation of children showing intolerance to artificial food colors may not be limited to those with ADHD.

(Kaplan et al, 1989)

Kaplan BJ, McNicol J, Conte RA and Moghadam HK. Dietary replacement in preschool-aged hyperactive boys. Pediatrics, 1989 (January), 83 (No.1): 7-17.

Kaplan et al (1989) designed a crossover dietary replacement study in hyperactive children to address several issues raised by previous research by including only preschool-aged boys with sleep problems and/or allergy-type somatic symptoms and notably using a replacement diet involving a broad range of food substances rather than targeting one class of substance. Twenty-four (24) pre-school male children (3.5 – 6 years) in day-care met the criteria of being diagnosed ADHD (during screening, children averaged more than 3 SD above the age-related mean on the parents Conners rating scale) and reported by parents to have sleep or physical problems. No stimulant medication was used during the study. The experimental diet was the Alberta Children's Hospital elimination diet (*ACH*) which eliminated artificial food colors, flavors, preservatives, MSG, chocolate, and caffeine, and decreased simple sugars; for some children the diet also eliminated milk and dairy products (15 children) and naturally occurring salicylates (4 children), for which the families reported possible problems. A placebo diet that matched the regular baseline diet was designed to appear *equivalent* to the experimental diet. Also, if the parent identified a food item (e.g., apples or carrots) to which the parent thought their child reacted, that item was eliminated from both the experimental ACH and placebo equivalent diet. Extensive efforts were carried out to ensure that the parents were effectively blind to treatment conditions and could not differentiate the placebo and experimental diets. The effectiveness of the blinding was verified in a post-study debriefing session with the parents. Procedurally, this study was conducted over a 10 week period consisting of an initial 3 weeks of baseline on regular diet, which was intended to determine normal eating patterns, to determine persistence of problem behaviors, and to obtain baseline data for outcome measures. This was followed by the dietary crossover phase of 3 weeks of placebo/equivalent diet and 4 weeks of ACH diet with these two diet treatments given in randomized counter-balanced order across subjects. All foods during dietary phase were prepared and delivered to families once each week. The outcome measures included: (1) <u>daily measures by parents</u>, including the Conners 10-item behavior rating with up to 4 additional individualized troublesome behavior items identified by parents, sleep measures, physical signs/symptoms, and food diary; (2) <u>daily measures by day-care workers</u>, including the Conners 10-item behavior rating; (3) a <u>planned additional daily observation by independent observers</u> which was terminated due to technical difficulties; and (4) <u>laboratory testing</u> conducted three times, once near the end of each test phase, consisting of <u>clinical</u>

measures/nutritional status (blood chemistry/hematology/basic physical) and planned psychometric testing (motor, sensory, attention, cognitive/learning/memory). It should be noted that the psychometric testing was not completed because of the investigators found the children to be "untestable" reportedly due in part to the nature of their ADHD behaviors and also to selection of inappropriate age-related tests. Statistical analyses of data from the last two weeks of each phase of study was conducted using repeated measures multivariate and univariate ANOVA with Tukey's method of multiple comparisons.

Study Results: The parent Conners behavior ratings showed a significant group effect for the ACH diet with the children's mean score (10.8) being significantly lower (i.e., improved) than their mean scores for either the placebo/equivalent diet (13.1) or baseline diet (14.5). Assessment of the parental data for individual children revealed that not all children were equally responsive to the ACH elimination diet but that 14 (68%) of the 24 children showed some behavioral improvement (9 showing at least a 25% improvement in behavior and 5 showing a milder improvement) on the ACH elimination diet as compared with the placebo equivalent diet. The remaining 10 children (42%) were unresponsive to dietary intervention. No placebo effects (baseline versus placebo/equivalent diet) or treatment x order effects were found. Analyses of the grouped data for the parent ratings of the individualized problem behaviors showed comparable significant improvement in behavior during the ACH diet compared with the placebo/equivalent diet (the number of individual responders was not identified). The parent sleep records also found that the ACH diet resulted in significantly shorter sleep latency (i.e., improved sleep measure) as compared with the placebo/equivalent diet treatment (the number of individual responders was not identified). However, the day-care worker daily ratings, using the same Conners 10-item behavior questionnaire that parents used, showed marginal but no confirmatory significant diet related differences in the children's behavior. Also, there were no reliable diet related effects for physical signs/symptoms or nutritional status. The incomplete psychometric testing precluded any meaningful statistical analysis of that data.

Study Assessment and Evaluation: In a group of 24 pre-school hyperactive children with sleep problems, parent ratings and recordings revealed a significant overall improvement in behavior and sleep when the children were on the ACH elimination diet compared with the placebo/equivalent diet. When the parent ratings data were viewed for each child, 9 children showed at least a 25% improvement in behavior and 5 children showed a milder improvement when being fed the ACH elimination diet compared with the placebo/equivalent diet. Although the investigators analyzed the ACH elimination data against both the baseline and placebo/equivalent diet data, the baseline diet is not the appropriate control for the ACH diet treatment. To assess diet related effects, the ACH elimination diet data should be compared with data from the proper diet control which is the placebo/equivalent diet. The remaining 10 children showed no diet related changes in the parent ratings. With the apparently effective procedures in this study for blinding the parents to treatment conditions, the parent-based detection of significant behavioral and sleep improvement in some of the children with the ACH elimination diet appears to be reliable. However, the day-care workers using the same behavior rating scale as the parents did not confirm a significant diet related change in behavior, although these day-care worker results were based on somewhat limited data. Unfortunately, the other two planned sources of behavioral outcome data, i.e. the psychometric testing and the behavior ratings by independent observers, were discontinued. Speculatively, since the investigators noted that psychometric testing could not be completed in part because of the children's ADHD behavior, this may suggest that the nature or scope of the effects associated with the ACH elimination diet apparently was not sufficiently effective in improving the children's behavior to enable their completion of the psychometric testing. But, the investigators did also note the difficulty in

selecting tests acceptable to this age-group of children. No diet related effects were found with any of the clinical or physical measurements conducted. Overall, since the day-care worker ratings, although completed on less than half of the subjects, did not confirm a significant diet related effect and since no other source of outcome behavioral data was available to resolve this inconsistency, the diet related findings in this study are considered inconclusive but suggestive that a special food elimination ACH diet may result in some level of behavioral and sleep improvement in some hyperactive pre-school children. While this suggests that some hyperactive children may be intolerant to some dietary elements, it is unknown what specific food item(s) or component(s) might be involved.

(Wilson and Scott, 1989)

Wilson N and Scott A. A double-blind assessment of additive intolerance in children using a 12 day challenge period at home. Clinical and Experimental Allergy, 1989, 19: 267-272.

Wilson and Scott (1989) designed a clinical double-blind challenge study to assess food additive intolerance under placebo controlled conditions. The study was conducted with 19 children (11 male, 8 female; 2-13 years) who were non-hyperactive but with a definite history of adverse reaction (respiratory, dermatologic, and/or behavioral) to artificial yellow color. (Note that 29 children were initially recruited and started on test but only 19 of these children completed the primary phases of this study) All subjects were reported to be responsive to a color-free diet and for most children color intolerance was verified in open (non-blind) challenge at home prior to the study. Four of the 19 subjects had a history that included adverse behavioral reaction to yellow color. The challenge items were administered in a bottled aide drink (250 ml of Lucozade with B-carotene coloring). Three types of drink, identical in appearance and flavor were used: (1) a color challenge with 8.5 mg tartrazine and 8.5 mg sunset yellow; (2) a preservative challenge with 12.5 mg sodium metabisulphite and 55 mg sodium benzoate; and (3) control/placebo aide. All drinks were reportedly identical in appearance and flavor but no procedures were described to verify the effectiveness of this blinding for parents or children. Procedurally, all children were started on an additive-free diet prior to the study and maintained on that diet throughout the period of testing. The initial study trial consisted of three 14-day phases with a different challenge drink for each phase. During each phase one bottle of the same type of drink was consumed daily at any time of day for 12 days. This was followed by a 2-day washout period after which the child was seen in the clinic (procedures in the clinic were not described). The three types of drinks were allocated double-blind in random order to each phase of testing. At the end of this initial study trial, the response to each type of drink was assessed with blind still maintained. If during one or two phases of that initial trial there was a suggestion that a deterioration in symptoms occurred, the child was invited to repeat the trial using the same protocol but with drinks recoded and the order randomized double-blind. A third trial (three 7-day phases with drinks taken for 6 days in each phase) was also offered if results were inconclusive. The outcome measures to assess treatment effects were admittedly more detailed and objective for asthma than for either dermatologic or behavioral effects. The assessment of behavioral responses was basically limited to two types of subjective unstructured assessments both by parents: (1) a daily diary of relevant symptoms with a simple scoring system ('daily maximum = 2'), and (2) an overall assessment at the end of each challenge phase with parents recording simply whether the child's symptoms were "better/worse/no difference". A deterioration in behavior was defined by a parent's assessment of "worse" confirmed by an increase in the total symptom score. When these differed, the result was considered

"inconclusive". The children were reported "seen in the clinic" after each challenge phase but no further description was given of whether or what clinical evaluations were conducted. The results were analyzed for individuals and not for the group as a whole. For each subject a total symptom score was calculated for each phase and a comparison was made between the three challenge phases of each trial. However, no statistical analyses of the data were apparently made. ***Study Results***: Out of 19 children, who were claimed to have a history of adverse effects from exposure to artificial yellow colors and definite benefit from the exclusion of artificial food colorings in their diets, only 3 showed any positive association between food color or preservative additives and adverse symptoms when challenged daily for 12 days under double-blind controlled conditions. Only 1 of these 3 "responders" developed adverse behavioral symptoms and that was in response to the preservative challenge but not to the color challenge. Interestingly, behavioral changes were not part of this child's prior history of adverse responses to artificial yellow color but this child had never previously been given food or drink with preservatives. The other two "responders" exhibited other symptoms without behavioral changes, one child developing urticaria when challenged with tartrazine/sunset yellow drinks and the other child developing intermittent asthma and abdominal pain with both the color and preservative challenge drinks. Notably, among the remaining16 children who showed no adverse effects of challenge under double-blind conditions were the 4 subjects who were claimed to have had a definite prior history of adverse behavioral responses to artificial yellow color. ***Study Assessment and Evaluation:*** None of the 19 children with a definite history of adverse effects from exposure to artificial yellow colors, including 4 with a history of adverse behavioral reactions, showed any adverse behavioral effects of challenge with combined tartrazine and sunset yellow under placebo controlled double-blind conditions. One child did exhibit an adverse behavioral response to the preservative challenge. Confidence in the study results, however, is somewhat limited in part by the questionable accuracy and sensitivity of the means used for detecting behavioral responses to the challenges. The assessment of behavior was totally subjective and unstructured with a very limited and poorly defined nonvalidated scoring system (behaviors scored as 1 or 2 and overall assessments as 'better, worse, no difference') and conducted only by the parents with no other source of assessment to confirm the parental observations. Although the investigators point out that the same criteria would have been used by the parents in the initial alleged association between additives and behavioral disturbance, this is hardly an appropriate basis for credible scientific assessment. Additional limitations in this study include no apparent verification of the effectiveness of the blinding for the parents and children (blinding of drinks with tartrazine is reportedly very difficult due to the intense coloration of the dye), and no description of statistical analyses of the data. Even considering these caveats, this study does suggest that anecdotal reports of definite adverse behavioral (and other) intolerance to artificial colors, specifically yellow colors, are not clearly evident when tested under more controlled clinical conditions.

(Pollock and Warner, 1990)

Pollock I, Warner JO. Effects of artificial food colours on childhood behaviour. Dis Child, 1990 (January), 65(1):74-7.

Pollock and Warner (1990) conducted a placebo controlled challenge study to investigate children with food additive intolerance. The study involved 19 children (1 hyperkinetic, 1 retarded; 3-15 years) with problem behaviors who were selected specifically based on parents' observations that their behaviors improved on a diet free of food/color additives and returned (primarily poor concentration and excess fidgetiness) within an average of 2.5 hours after

consuming food additives. According to the parents, even small amounts of food additives, precipitated the recurrence of behavioral problems. (Note that 39 children were recruited for this study but only 19 completed all testing.) The challenge item was a gelatin capsule with/without 125 mg blend of artificial food colors (50 mg tartrazine, 25 mg sunset yellow, 25 mg carmoisine, 25 mg amaranth). The placebo contained lactose. The gelatin capsules were made opaque with iron oxide which, according to the investigators, effectively disguised the color additive capsules so that they could not be distinguished from the placebo capsules. The parents were told only that the capsules for each week contained either food additives or placebo, but parents were not asked whether they could differentiate the color and placebo capsules. Procedurally, all children were maintained on their food additive elimination diets throughout the test period. The 7 week double blind placebo controlled challenge study consisted of 2 separate weeks of color challenge and 5 weeks of placebo with 3 placebo weeks between each color challenge week. Children were randomly assigned to one of two treatment sequences (C P P P C P P or P P C P P P C) and received an single
appropriate capsule (color or placebo) daily. Two outcome measures, used to assess treatment effects, were both provided by the parents and consisted of: (1) a daily rating of the child's behavioral and somatic symptoms (Conners Hyperactivity Index supplemented with questions about symptoms associated with allergies) and (2) a weekly overall behavior assessment (recorded as improved/same/worse). Data were analyzed using the Wilcoxin rank tests.

Study Results: The parent daily behavior ratings were significantly higher (enhanced problem behaviors) for the color challenge weeks than for placebo weeks and there was no confounding treatment order effects. The daily somatic scores (associated with allergies) showed no differences between color and placebo challenge weeks, indicating that the changes in behavior scores were not related to or secondary to somatic symptoms. Additional analyses showed no carry-over of the color challenge effect into the washout week and also showed no difference in behavior ratings between week days 1 and 7 indicating that the behavioral changes occurred early in the challenge week and that the effects of food color intolerance do not result from a cumulative effect of food colors. Notably, however, the absence of any differences between color and placebo weeks in the parents' overall assessment of behavior at the end of each week provides no supportive confirmation for the color effects seen with the parent daily behavioral ratings.

Study Assessment and Evaluation: This study reported that a daily intake of 125 mg mixed artificial colors for 1 week, relative to placebo, can produce a small but measurable change in the parents daily rating showing an increase in problem behaviors for non-ADHD children. This effect was unrelated to the somatic (allergy) symptoms or atopic state of the children. However, the reliability of these findings, or at least their biological relevance, appear to be questionable due primarily to the fact that most of these parents were unable to confirm a color related worsening of behavior in these children in their weekly overall behavioral assessments. Unfortunately, the use of only parents' observation of behavior to measure treatment related effects does not enable independent sources of information to help resolve this inconsistency. While the study provided no verification of the effectiveness of the blinding procedures for the parents, the investigators do suggest that any ineffective blinding of the parents seems unlikely since their weekly behavioral assessments might then have more accurately reflected the daily behavior ratings. Overall, since the claim by parents of being able to detect deterioration of behavioral after even consumption of small amounts of food additives was a criterion for entry into the study, the fact that the parents weekly assessment could not detect an overall behavioral change from daily challenge to 125 mg artificial colors, a rather high dosage, for a week (on two separate week trials) and the rather small change in behavior detected with the parents' daily

Connors assessment indicate that there is little substantial or functionally (biologically) relevant adverse effects of artificial colors on the behavior of children anecdotally identified as having some food intolerance.

(Sarantinos et al, 1990)

Sarantinos J, Rowe KS and Briggs DR. Synthetic Food Colouring and Behavioral Change In Children With Attention Deficit Disorder: A Double-Blind, Placebo Controlled (Challenge Study). Proc. Nutr. Aust., 1990, 15: 233.

(It should be noted that the review and evaluation of this clinical study will be limited since it is based only on an Abstract Summary of the study which contains few details of the experimental design, test group assignments, and test results. A full published report of this study was not available. Consequently, only a tentative assessment and evaluation of the study and results can be presented.)

Sarantinos et al (1990) conducted a placebo controlled challenge study to investigate the effects of two artificial food colors, tartrazine and sunset yellow, on the behavior of children diagnosed with attention deficit disorder (ADD). The investigators recruited 13 children (12 male, 1 female; 4-14 years) diagnosed with ADD and previously placed on a diet free of artificial food colors. Four to six children were reportedly intolerant to various foods, two of which were behaviorally intolerant to orange juice (the challenge vehicle used in this study). Nine parents considered their child improved on the diet, while 4 parents were uncertain about improvement. The challenge item was orange juice with/without 10 mg tartrazine or 10 mg sunset yellow. The orange juice drinks were consumed from an opaque pack using a straw. No procedures for determining or validating the effectiveness of the blind were described. All children were placed on color-free diet at least 6 weeks prior to study and remained on color-free diet throughout the 28-day study. Procedurally, the 13 children were assigned to one of two groups (report did not state whether assignment was randomized and did not specify distribution of subjects or composition of the two groups). Over the 28-day study period, the children in Group 1 (n=6 or 7) were challenged only with tartrazine on 6 random days (not specified whether all children were given color challenge on same days). The children Group 2 (n=6 or 7) were also given color challenges on 6 random days, tartrazine on 3 of those days and sunset yellow on the 3 other days. A 2-3 day washout (placebo) period was allowed between each challenge day. On all non-color challenge days, children received a placebo drink. As the outcome measure, behavior was assessed daily by parents only using both the Conners' Abbreviated Parental Scale and the Rowe Behavioural Rating Inventory Scale. Data were analyzed with repeated measures analysis of variance.

Study Results: Only 12 subjects completed the double-blind challenge trial but the report did not specify which one child was excluded. Both parental rating scales showed that 2 of the 6 or 7 ADD subjects in Test Group 2 showed significant adverse change in behavior (i.e., increased irritability, impulsivity, restlessness and sleep disturbance) associated with both color challenges (tartrazine and sunset yellow). Both children were atopic and among those whom parents had reported to be responsive to color-free diets; one child was intolerant to multiple food items and the mother could correctly identify (25/28 times) whether the child received color or placebo (unclear whether guess was based on behavior or broken blind).Apparently, none of the 6 or 7 children in Group 1 responded to the challenge with tartrazine alone.

Study Assessment and Evaluation: Due to the absence of a full report for this study, there are numerous uncertainties regarding the study design, conduct and results. As a result, there are several aspects of this study as described in the Abstract Summary of this study which minimize

confidence in the reliability of this study's findings and confound interpretation of the study results. These include, for example: (1) 13 children were recruited for the study but only 6-7 children comprise each of the two treatment groups; (2) unknown composition of the two treatment groups, e.g., how many diet responders were in Group 1 and Group 2; (3) the questionable validity of using 2 subjects who were reported behaviorally intolerant to orange juice for this study in which orange juice was the challenge vehicle and placebo treatment; (4) the use of only parents to assess treatment effects without any other source of behavior assessment to confirm the parents' reported treatment effects; (5) the unclear effectiveness of the blinding for parents and children; (6) no actual data are presented in this abstract summary; and (7) questionable significance of no treatment effects in Group 1 with tartrazine challenge but significant treatment effects in Group 2 children who were also given tartrazine challenge randomly alternated on separate days with sunset yellow challenge. Overall, based on the information available, the absence of other behavioral measures to corroborate the changes in parental ratings and the uncertainties engendered in the absence of various details about the experimental design, the effectiveness of the blinding conditions and data analysis certainly raise questions about the confidence in reliability and biological significance of the findings, precluding any credible interpretation of these findings with regard to the possible association of the artificial colors, tartrazine and sunset yellow, and adverse behavioral effects in children with ADD.

(Egger et al, 1992)

Egger J., Stolla A and McEwen LM. Controlled trial of hyposensitization in children with food-induced hyperkinetic syndrome. The Lancet, 1992 (May), 339: 1150-1153.

Based on the premise, suggested by previous research (see reviews of Egger et al, 1985; Kaplan et al, 1989), that food intolerance is a possible cause of the hyperkinetic syndrome in some children and may involve some allergic/immunologic mechanism, Egger et al (1992) designed a placebo controlled clinical study to determine whether hyperkinetic children can be desensitized to the adverse effects of provoking foods using an enzyme-potentiated desensitization treatment (EPD). This study was conducted in four phases: Phase I – identify hyperkinetic children responsive to an open (non-blind) oligoantigenic diet (i.e., a very restrictive diet that eliminates every kind of food and food additive assumed to provoke adverse behavioral reactions in certain children); Phase II – non-blinded identification of specific provoking food items; Phase III – double blind, placebo controlled EPD treatment; and Phase IV – determine effectiveness of treatment by challenge with provoking food items. In Phase I, 185 children diagnosed with hyperkinetic syndrome (Conners score > 15) received an open (non-blind) oligoantigenic diet for 4 weeks. Based on parent and teacher non-blinded observations and a Conners score <15, 116 children (63%) reportedly improved on the diet and were entered into the food reintroduction phase of study. For Phase II, all 116 food intolerant children were maintained on their oligoantigenic diet and food items eliminated from that diet were sequentially reintroduced under non-blind conditions singly every 5 days. Identification of provoking (intolerant) food items was based on recurrence of hyperkinetic behavior (overactivity) or other associated symptoms (e.g., headache and recurrent abdominal symptoms such as discomfort, bloating, diarrhea) on at least three trials with that food item. In the experimental treatment Phase III, 40 of the food-intolerant hyperkinetic children (36 male, 4 female; 3-15 years/ average 9 years) agreed to take part in the double-blind, placebo-controlled trials. All children were maintained on an oligoantigenic diet and provoking foods were avoided during this period. Subjects were randomized to receive 3

intradermal injections of either placebo/buffer solution (n=20; Treatment Group 1) or EPD treatment solution (n=20; Treatment Group 2) at two-monthly intervals. The EPD treatment included mixed food antigens (multiple foods, additives, colors and preservatives). Only one principle investigator knew the key to the randomized EPD/Placebo treatments. Children and parents were not told which treatment was administered. The investigators stated that "placebo and active [EPD] treatments were both colourless solutions", ostensibly blinding the children and parents to any apparent difference between the two treatment solutions (see comment regarding blinding of treatments in Study Assessment and Evaluation, below). The Phase IV assessment of the effectiveness of treatment was initiated three weeks after the third EPD injection, when provoking foods for each subject were reintroduced, one item at a time, while the children were still maintained on their oligoantigenic diets. For evaluating responses to the various provoking food items, the parents were asked to maintain a daily diary card of hyperactivity and other symptoms (e.g., headache and recurrent abdominal symptoms such as discomfort, bloating, diarrhea). As one measure of treatment effectiveness, parents were told to stop the food when food-related symptoms occurred and persisted for 24 hours. As a second measure of treatment effectiveness, at the end of introducing all provoking foods, parents were asked whether the treatment had been successful in preventing or reducing symptoms on eating one or more of the provoking foods. Statistical analyses (Fisher's exact test) were conducted for (1) parents' stopping the first provoking food given to the EPD and placebo groups, and (2) response to whether parents considered treatment successful. Skin tests were also conducted to identify atopic children and serum IgE was measured.

Study Results: A variety of food items were identified as provoking behavioral and associated symptoms; chocolate, artificial colors, cow milk, egg, citrus, wheat, beet sugar and nuts were among the most common provoking food items. Adverse response to provoking food items (based on analysis of the number of parents stopping the first provoking food reintroduction) occurred less frequently in the EPD treated group than in the placebo treated group. Specifically, 15 of the 20 children given EPD treatment were able to eat previously identified provoking food items without adverse reactions, but only 7 out of 20 placebo treated children were able to continue eating provoking foods. Also, after all provoking food items were reintroduced, more parents of children treated with EPD than those treated with placebo, thought the treatment was successful. While these results depended on recurrence of any adverse responses, behavioral or associated symptoms, to the provoking food items, the abdominal symptoms, such as discomfort, bloating and diarrhea, were usually the first to appear (it is assumed but not specifically stated that hyperactive behaviors occurred subsequently, as well). Reaction to provoking foods tended to recur several months after completion of the study but tolerance to the provoking foods could be reinstated with additional EPD treatment. At the end of the trial, 16 of the children in the placebo/Group 1 were given EPD treatment and 15 developed tolerance to previously provoking foods (it was unclear whether this was conducted under non-blinded conditions). These study results were not related to atopic status or IgE levels in these children.

Study Assessment and Evaluation: This study provided some information about the proportion of hyperkinetic (hyperactive) children that may react adversely to some foods. Based on observations by parents and teachers under non-blinded conditions, approximately 63% (116/185) of children diagnosed with hyperkinetic syndrome appeared to be food intolerant, showing positive response to an oligoantigenic diet which eliminated provoking food items and showing adverse response to reintroduction of those provoking food items. When 40 of these apparent food intolerant hyperkinetic children were tested under double-blind, placebo controlled conditions to determine whether they could be desensitized (EPD treatment) to the adverse effects of the provoking food items, it was shown that more of the EPD treated children (15/20)

were able to eat previously identified provoking food items than placebo treated children (7/20). Several issues involving experimental design weaknesses affect the reliability and biological relevance of this study's finding: (1) One particular issue of concern involves the adequacy of the blinding of the experimental treatment which is central to the reliability of the subjective parental observations. The investigators stated that "placebo and active [EPD] treatments were both colourless solutions". The placebo was simply a colorless buffer solution. However, it is difficult to understand how the active (EPD) solution could have been colorless in view of the fact that the mixed antigens in that preparation included multiple food colors (tartrazine, chocolate brown, Ponceau 3R, erythrosine, green S, and Annatto). Any notable difference in coloration between the placebo and EPD solutions would make blinding of the parents and children to the treatment very difficult and introduces the possibility of unintentional confounding assessment bias. (2) A second study limitation which affects confidence in the reliability of the study findings, or at least their biological relevance, involves the fact that the parents were the only source for detecting response to provoking foods and assessing effectiveness of the desensitization treatment. In the absence of supporting data from sources other than parents, there was no corroboration of the parental based findings. And, a third issues involves the fact that (3) A final study issue involves the minimal assessment of specific behavioral changes during the double-blind, placebo controlled testing of the effectiveness of desensitization to the adverse effects of provoking food items. The effectiveness of treatment (placebo and EPD) was based on the parents' determination of whether reintroduction of foods should be stopped and on the end of trial parents' overall assessment of improvement in symptoms. Responses to the various provoking food items were evaluated by parents who were asked to maintain a daily diary card of hyperactivity and other symptoms, but no structured or validated rating system, e.g. Conners hyperactivity rating scale, was used to help characterize the scope of behavioral changes, even though Conners behavior ratings were conducted by parents in the initial non-blinded identification of responders to the oligoantigenic diet. The investigators do not even provide data identifying which food-related behavioral change or associated symptoms were used by the parents to decide whether the provoking food should be stopped. In fact no behavioral data are presented or analyzed to show whether any food-related behavioral changes even occurred during the re-introduction of provoking foods. And, beyond the statement that abdominal symptoms were typically the earliest signs of adverse response to provoking food items, no data for those associated symptoms were presented. Overall, the questionable confidence in the reliability of the study data and the minimal specific assessment of behavioral changes due to provoking food items limit the potential significance and interpretability of this study's findings relating food and hyperactive behaviors. In view of the caveats inherent in these study weaknesses the results of this study may be viewed only as suggestive that intolerance to a variety of food items may be causally associated with hyperkinetic behavior and symptoms in some children and that some food intolerant hyperactive children may be desensitized to the adverse effects of provoking food items.

(Carter et al 1993)

Carter CM, Urbanowicz M, Hemsley R, Mantilla L, Strobel S, Graham PJ, Taylor E. Effects of a few food diet in attention deficit disorder. Arch Dis Child, 1993 (November), 69(5):564-8.

A previous study suggested that foods and additives could affect hyperactive behavior adversely (Egger et al, 1985), based on the experimental design from that study Carter et al (1993)

conducted a multiphase study of the effects of a restricted food elimination ("few foods") diet on ADHD children. The "few foods" diet is virtually identical to the oligoantigenic diet (Egger et al, 1985) both of which are very restrictive diets that eliminate every kind of food and food additive assumed to provoke behavioral reactions in certain children. The main purpose of the study was to determine whether anecdotal reports of food intolerance in ADHD children may be due to a 'placebo effect' of expectation and suggestion. The study was conducted in three phases: Phase I – identify ADHD who are food intolerant and improve behaviorally on a "few foods" diet; Phase II – open (non-blind) reintroduction of food items and additives to identify food components responsible for adverse reactions; and Phase III – experimental double-blind, placebo controlled re-introduction of incriminated foods and additives. Seventy-eight (78) ADHD children (69 male, 9 female; 3 – 12 years) were used for Phase I testing. Those children who were already on some special diet (43) were asked to stop those diets at least one week prior to this study. All children were then given an open (non-blind) restricted "few foods" diet for three to four weeks. The foods allowed were typically two meats, two carbohydrate sources, several root and green vegetables, bottle water, sunflower oil, and milk free margarine; for individual children the diet also avoided foods already suspected. Approximately 76% (59/78) of the ADHD children were found to be food intolerant based on their reported behavioral improvement on the diet. Forty-seven (47) of these "diet responders" were used for Phase II testing in which the "few foods" diet was continued and various foods and additives were openly reintroduced (non-blinded) at the rate of one per week to tentatively identify those that may elicit an adverse response. Food additives were usually given initially as mixtures in additive-containing foods. When adverse reactions occurred and colors were suspected, response to artificial colors was specifically tested using capsules with 6.5 mg mixed colors. A large number of foods were implicated during the open reintroduction phase. Most reactions to food included worsening of behavior sometimes accompanied by physical symptoms. For the Phase III double-blind placebo controlled challenge with provoking food items, 19 diet responsive ADHD children were used (23 children were started in Phase III but 4 did not complete testing). The active food challenges consisted of multiple food items, including artificial colors and preservatives, incriminated as provoking adverse behavioral responses in Phase II. Each provoking food item was selected to be adequately disguised in a food known to be tolerated (to act a excipient and placebo). Artificial colors were generally given in capsules rendered opaque with iron oxide with each capsule containing 6.5 mg of mixed colors; the number of capsules given varied, but no child received more than 4 capsules or 26 mg daily. Glucose capsules were used a placebo. Only the study director, who took no part in actual testing, knew the treatment code and everyone else remained blind to the order in which active and placebo foods were given. To ensure that the taste and appearance of the placebo and active food treatments were indistinguishable, on the first day of each test week the families were asked to guess whether they had been given the active or placebo food item; none were able to guess correctly beyond chance, verifying the effectiveness of the blind. Procedurally, all children were maintained on their "few foods" diet throughout the Phase III period of testing. Each incriminated food item or additive and its appropriate placebo were given daily each for one week (experimental period) with a two-week washout period between active and placebo treatments. The order in which incriminated foods and placebo were given was allocated randomly. Most of the children received more than one provoking food item. The outcome measures to assess behavior were taken at three time points: at entry to the experimental Phase III, at the end of each experimental week (placebo or food item). These measures included: (1) parent behavior ratings (Conners parent rating scale) and parent global rating of severity of behavior problems; (2) psychologist rating of hyperactive behavior (inattentiveness, restlessness, fidgetiness); (3) paired associate learning test; and (4) matching

familiar figures with a simplified form of testing for children under 6 years (impulsiveness). Most data were statistically analyzed using t-tests. The parent global ratings were tested by Fisher's exact probability test.

Test Results: Approximately 76% (59/78) of the ADHD children were found to be food intolerant based on their improvement on an open (non-blind) 'few foods' elimination diet. In the double blind phase of testing, parent ratings of behavioral change showed statistically significant deterioration of behavior following the incriminated food item challenge compared with the placebo, but the magnitude of this behavioral effect was shown to be rather small (Conners score of 8.8 for placebo and 13.9 for food item challenge which did not reach the conventional hyperactivity cutoff score of 15) and the types of behaviors rated by the parents as most affected suggested a possibly greater effect on irritability than on attention deficit. (The parent global ratings were described as favoring the placebo and showing fewer treatment failures. But, none of these data were presented and the actual meaning of these stated effects was not interpretable.) The psychologist ratings also showed significantly more hyperactive behavior, primarily fidgetiness, following the food item challenge than placebo, as did the matching familiar figures test of impulsiveness. Paired associate learning, however, was unaffected.

Study Assessment and Evaluation: Both parent and psychologist ratings and one of two objective psychological tests (impulsiveness) showed significant behavioral changes associated with food item challenges for the entire group of subjects. While these behavioral changes were generally in the direction of negative (worsening) effects, the small magnitude of these changes makes their biological significance unclear. Possibly, combined exposure to multiple provoking food items may be necessary to elicit an additive effect of a greater magnitude behavioral response. The possibility of additive effects should be addressed in subsequent studies. Notably, the types of behaviors rated by the parents showed a possibly greater effect on irritability than on attention deficit, suggesting behavioral changes not necessarily associated only with the hyperactivity syndrome. Unfortunately, the data for individual subjects were not assessed or presented to determine whether some of the children may have exhibited stronger responses and behavioral changes possibly more closely associated with attention deficits. Also, although most of the children did receive more than one provoking food item, it was not stated whether those children exhibited adverse responses to all challenge food items and whether those responses were similar in type and magnitude. In assessing the overall significance of these study findings two minor study weaknesses should be mentioned. First, even though food challenge was given for a whole week, families were asked to guess if they had been given the provocative or placebo food only on the first day of each challenge week. While there may be some uncertainty as to whether the effectiveness of the parents' blind was maintained throughout the test week, a reasonable effort was made verifying the effective blinding of parents to active food and placebo treatments. Second, while the use of uncorrected multiple t-testing may have increased the probability of false positives (i.e., erroneous findings), it is unlikely that many, if any, of the treatment effects in this study were false positives. These study weaknesses are considered minor with little impact on the reliability of the study findings of small but significant adverse behavioral effects of a variety of provoking food items, which were confirmed across several sources of measurement (parent, psychologist, objective testing). Overall, these findings do show that parents' reports of behavior change with a "few foods" elimination diet can be confirmed to some extent, in a selected group of ADHD children, by double blind, placebo controlled trial with multiple sources of valid testing. These findings also support the conclusion that some ADHD children have an intolerance to a variety of food items, including but not limited to artificial food colors and additives, which may cause or exaggerate adverse problem behaviors but which may or may not be related to the hyperactivity syndrome. However, the use of general

food items in this study makes it difficult to identify specific food chemicals that may be causing the adverse effects.

(Boris et al, 1994)

Boris M and Mandel FS. Foods and additives are common causes of the attention deficit hyperactive disorder in children. Ann Allergy, 1994 (May), 72(5):462-8.

The stated purpose of this study was to determine whether dietary components play a role in attention deficit hyperactivity disorder (ADHD) in children. ADHD children showing improved behavior on an open (non-blind) multiple food elimination diet (Phase I) and responding adversely to certain food components in an open challenge (Phase II), were subsequently tested in a double blind placebo controlled food challenge (DBPCFC) to determine whether the responses to suspect provoking food components were reproducible under controlled test conditions (Phase III).

Phase I - Twenty-six children with ADHD were placed on an open (i.e. not blinded) multiple food elimination diet for 2 weeks to identify children that appear to show behavioral improvement on the elimination diet. This diet eliminated any components that consisted of dairy products, wheat, corn, yeast, soy, citrus, egg, chocolate, peanuts, and artificial colors and preservatives. Prior to and at the conclusion of the 2 week elimination diet phase, the parents completed a Conners Parent Rating Scale-48 (CPRS-48) which included a subset of questions used in this study as the quantitative index of hyperactivity. A score of 15 or greater was taken to identify hyperactive children. Nineteen or 73% of the 26 children appeared to show improved behavior (mean hyperactivity score of 9.6) on the open elimination diet, as compared with their original diets (mean hyperactivity score of 25). Significantly more of the 19 “diet responders” were atopic than of the 7 “non-diet responders”.

Phase II - The 19 "diet responders" were then given open (i.e. not blinded) food challenges over the next month. The parents challenged their child with one restricted diet item every two days; reactions to a diet item were retested to confirm the reaction. All 19 children reacted to multiple diet items during the open challenge.

Phase III – The 19 “diet responders” were then enrolled in a 7-day DBPCFC trial, but only 16 completed the entire trial (11 male, 5 female; 7.5 years). Each set of parents were asked to select the most reactive food or agent that appeared to induce the recurrence of symptoms in their child during open challenge (Phase II) and that one item was used to challenge that child during the DBPCFC. Challenge items (number and sex of subjects) included the following: milk (2m, 3f), colors (4m), corn (3m), wheat (2f), soy (1m), and oranges (1m). Note that none of the subjects responded to preservatives in open challenge. The challenge material consisted of5 grams of the powdered food or 100 mg of a blend of food colors. To control for the taste and color of the challenge items, either lentil soup or an apple cranberry sauce was used as the excipient and as the placebo vehicle. Seven dieticians tasted the vehicles with each of the challenge items and could not distinguish a difference in taste. However, the parents and children were not asked to verify whether they could differentiate the placebo from the challenge item. Procedurally, all children were continued on the elimination diet throughout double-blind testing. Over the 7-day challenge period, days 1, 2, and 7 were designated placebo days for all subjects. On days 3, 4, 5 and 6, the food challenge and placebo treatments were randomly administered with 13 subjects receiving food challenge two times, 2 subjects receiving food challenge three times, and 1 subject receiving food challenge one time. Both the parents and primary investigators were blinded to the challenge order during this 7-day period. The study was carried out in the home environment during a school vacation. As the sole outcome measure, the parents completed a

Conners Parent Rating Scale-48 (CPRS-48) which included a subset of 10 questions used in this study as the quantitative index of hyperactivity. The parent ratings were completed prior to the study while children were still on their original diets, at the end of the open elimination diet period, and daily during the 7-day challenge period. Statistical analyses of behavioral data were conducted primarily using the repeated measures analysis of variance and repeated measures analysis of covariance..

Study Results: Nineteen or 73% of the 26 ADHD children given an open (non-blind) elimination diet appeared to show marked behavioral improvement, as compared with their original diets (mean hyperactivity scores of 9.6 and 25, respectively). A significantly greater number of atopic subjects (i.e., history of allergic manifestations) than non-atopic subjects were among the "diet responders" (15 atopics and 4 non-atopics responded to the elimination diet). The 16 "diet responders" that completed the double-blind, placebo controlled challenge trial collectively showed significantly elevated hyperactivity scores with the food challenges (mean Conners score of 18.1)as compared with the placebo treatments (mean Conners score of 8.2). However, inspection of the graphed Conners scores showed that only 11 of the 16 subjects actually responded to the active food challenges. Five (5) subjects did not respond to their active food challenge under blinded conditions even though they were reported to show improved behaviors when switched from their original diet to the open elimination diet and to show marked adverse behavioral effects when given provoking food items in open (non-blinded) challenge. In the double blind challenge, the authors do not identify what food items these 5 non-responding subjects were receiving. Nor is it know whether the 4 children that received the artificial colors as their food challenge did or did not respond. No information was presented as to whether the elimination diet or the exposure to provoking food items affected the atopy/allergic status of the children

Study Assessment and Evaluation: In open (non-blind) diet and challenge trials, 19 or 73% of 26 ADHD children were reported to be food intolerant, showing behavioral improvement with a multiple foods elimination diet and adverse behavioral changes elicited by various food components (reacting to at least three or more different food items). Interestingly, in this study more atopic children than non-atopics responded to the elimination diet, contradicting an earlier study which found no involvement of atopy (Egger et al, 1985). Sixteen of these "diet responders" were subsequently tested in a double blind, placebo controlled challenge trial and 11 children were confirmed to be food intolerant. Based on parental ratings, these children exhibited adverse behavioral responses (elevated hyperactivity scores) to challenge with previously identified provocative food items; the other 5 children did not react to the food challenges. The investigators also noted a substantial difference in hyperactivity scores between the original (pre-study) diet period and the active food challenge days (mean Conners hyperactivity scores of 25 versus 18.1, respectively). They suggested that this showed multiple food diets to have a greater adverse effect on the behavior of hyperactive children than a single item food challenge based on the fact that only one provoking food was used in the challenge but the children were reactive to multiple provoking foods in their regular diet. While conceptually it seems plausible that exposure to multiple provoking food items might be expected to have a greater (additive) adverse affect than challenge with a single provoking food item, in this instance, however, the difference between the two scores of hyperactivity may be more apparent than real since the conditions under which the two ratings were conducted were completely different. The parental rating of the hyperactive behaviors prior to the study (with the regular diets) was conducted under non-blinded conditions (i.e., they knew that their child's hyperactivity was not being treated in any manner) and, consequently, they may have had a biased perception of their child exhibiting a greater level of hyperactive behavior. However, when the parents rated the behaviors after a

blinded challenge treatment, their uncertainty may have resulted in a more conservative (lower score) assessment of their child's behavior.

As noted previously, the significant food related challenge effects in this study were based on parental ratings (Conners hyperactivity index) as the only measure of behavioral response. However, in the absence of other confirmatory outcome measures, confidence in these findings is limited. A more reliable assessment of the relevance of the parental findings could be made if other measures of behavior were included, e.g. clinical evaluation and teacher evaluation. Furthermore, although the parents and primary investigators were blinded to the challenge order during the testing period and several dieticians were unable to distinguish the taste of the active food item versus the placebo treatments, there is still an uncertainty about the adequacy of the study blind, since no specific effort was made to determine whether parents/children were able to distinguish the two treatment items. Verification of the effectiveness of the blinding conditions, particularly in a study that uses subjective parental evaluation as the sole measure of treatment related behavioral effects, is critical for the integrity and reliability of any findings in a controlled challenge study. Overall, in view of the uncertainty about the effectiveness of the blinding for the parents and children and particularly in conjunction with the fact that no other measures of behavioral change were used to verify the parental reports of adverse behavioral response to provocative food challenges, the confidence in the reliability of the study findings showing adverse behavioral effects of challenge with various provoking food items is limited and the relevance of these observed behavioral changes is unclear. The findings may at least be considered suggestive that intolerance to various food components, possibly including but not limited to artificial colors, may cause or exaggerate hyperactive behaviors in some ADHD children. While the use of general food items in this study makes it difficult to even tentatively identify specifically the food elements (component) that are causing the adverse effects, none of the subjects were reported to have responded to preservatives in open challenge.

With regard to the 73% of ADHD children who responded to the open elimination diet, the present study also showed that ADHD children who were atopic (history of allergic symptoms) were more likely to respond to the elimination diet than those children who were non-atopic, suggesting a possible association between allergy/immune function and reported food intolerance in ADHD children. This contradicted the results from another study (Egger at al, 1985) in which atopy was equal in diet responders and non-responders. Although the present study did not report whether the subjects exhibited allergic reactions to the challenge food items, one possibility suggested by these non-blind observations is that the reportedly responsive subjects were somehow immunologically sensitive to the provoking food items and that the negative behavioral changes were secondary to or an associated consequence of the "allergic" or immunological response. In lieu of an "allergic"/immunological role, the investigators suggest the possibility that the challenge items could be having a more direct adverse effect on the nervous system. Based on the information presented in this study it is not possible to determine which of these alternative interpretations may be more feasible.

(Rowe and Rowe, 1994)

Rowe KS, Rowe KJ. Synthetic food coloring and behavior: a dose response effect in a double-blind, placebo-controlled, repeated-measures study. J Pediatr, 1994 (November), 125(5 Pt 1):691-8

The objective of this challenge study was to establish whether there is an association between

tartrazine, an artificial food color, and behavioral change in children referred for assessment of suspected hyperactivity. Uniquely, a full dose response assessment of tartrazine was used as the study design. The first stage of this study identified children who might be sensitive to artificial colors. For this, 200 children with problem behaviors (suspected but not diagnosed Hyperactive) were given a 6-week open trial of a diet free of artificial colors. The parents of 150 or 75% of the 200 reported behavioral improvement with the diet but deterioration in behavior with ingestion of foods containing artificial colors. The primary behaviors consistently reported related to irritability, restlessness, and sleep disturbance (behaviors not typically characteristic of hyperactivity syndrome). Thirty-four (34) of these "diet responders" and 20 children without problem behaviors given the color-free diet for 6 weeks participated in the double-blind, placebo controlled, repeated measures challenge study (38 male, 16 female; 2-14 years @ average 7 years). The "diet responders" were categorized by parents as likely diet responders (23) or uncertain diet responders (11). For children ≥6 years of age the challenge item was a double capsule containing either lactose (placebo) or one of six dosages of tartrazine (yellow 5) in the inner capsule surrounded by lactose. For children <6 years the challenge item was packaged orange juice with a straw (250 ml) with/without one of six dosages of tartrazine. The level of tartrazine used were: 1, 2, 5, 10, 20, and 50 mg (the latter two highest dosages were administered toward the end of the test period). Other than use of specially designed capsules, no description was given as to blinding procedures or to verification of effectiveness of blinding. Procedurally, all children were maintained on a color-free diet and given one capsule (or orange juice) on each morning of the 21-day study period. For each child the tartrazine dosages were administered randomly once each on a different day, except that the 20 and 50 mg dosages were administered toward the end of the study period. The design allowed for a placebo lead-in period of 3 days and a placebo washout period of at least 2 days between each day of tartrazine challenge for each child. The outcome measures for behavioral assessment consisted of two ratings scales completed by the parents daily throughout the study at the end of each 24-hour period to ensure that the sleep period was included. The two rating scales were: (1) a 30-item behavioral rating inventory, which was devised by the investigators based on a composite analysis of the clinical histories of 50 suspect reactors and was validated to discriminate between color and placebo ingestion, and (2) the Conners 10-item parent/teacher questionnaire, included for comparative purposes only (none of the Conners rating data were statistically analyzed for treatment effects). Atopic histories and allergic reactions to allergens and tartrazine were also obtained. Standard parametric and non-parametric statistical techniques were used to analyze the data.

Study Results: In the initial stage of this study, 150 or 75% of the 200 children with problem behaviors (suspected but not diagnosed Hyperactive) were reported by parents to be adversely responsive to foods with artificial colors, showing primarily increased irritability, restlessness, and sleep disturbance. Under double-blind challenge conditions, the parent behavioral inventory ratings identified a total of 24 of the 54 children tested as showing consistent behavioral reaction to the tartrazine challenges in clearly dose related fashion with amplitude and duration of effect increasing with increasing dosage level. Beyond 10 mg there appeared to be a ceiling effect in magnitude but higher doses increased duration of effect. The tartrazine challenge reactors included 19 of the 23 "likely diet responders", 3 of the 11 "uncertain diet responders", and 2 of the 20 control subjects. When the mean parent behavioral ratings were used to analyze the differences between placebo and tartrazine challenge, significant differences were found at all six dosage levels. There were no sex related (male/female) or age related (2-6 years/7-14 years) differences between the tartrazine reactors and non-reactors. All tartrazine reactors, including the 2 control subjects, were atopic but none showed an allergic reaction to tartrazine. However, it was not stated whether the non-reactors were atopic or had an allergic reaction to tartrazine. The

behavioral changes were independent of tartrazine's known allergenic effects.

Study Assessment and Evaluation:

This parental behavioral ratings in this study revealed a dose related functional relationship between the artificial food color tartrazine (Yellow 5) and adverse behavioral changes in 24 of 54 children at dosage levels of tartrazine ranging from 1 mg to 50 mg. These tartrazine reactors included 19 of 23 diet-responsive children with behavioral problems (suspected but not diagnosed hyperactive), 3 of 11 uncertain diet-responsive children also with problem behaviors (suspected but not diagnosed hyperactive), and 2 of 20 control subjects (not diet responsive and no problem behaviors). There were no sex-related or age-related (preschool versus school aged) differences in these tartrazine effects. The investigators reported that the main behavioral features associated with the double-blind tartrazine exposure, as rated by parents using the devised Behavioral Rating Inventory, were irritability, restlessness, and sleep disturbance, but apparently not behaviors associated with hyperactivity such as inattention. Tabulations of these specific rating data were not presented in the study report. It is possible that the devised rating system itself may have inadvertently favored reporting of those three behaviors over the reporting of attention related behaviors, since there were more rating elements for irritability (11), sleep disturbance (9) and restlessness (4), than for attention span (3). Some suggestion that attention may have been affected to some extent by the tartrazine treatment is provided by the graphic presentations of rating data for two reactive subjects based on the Conners hyperactivity index , in which attention behavior is a core element, showing apparent dose related effects of tartrazine. Unfortunately, the parental ratings with the Conners index were not tabulated either in this report. Since the Behavioral Rating Inventory was not properly validated and without presentation and analysis of all of these rating elements, the question of whether treatment behaviors related to inattention may have occurred but were inadvertently underestimated cannot be resolved.

There are two primary experimental issues which could potentially affect the confidence in the reliability of the basic findings from this study, one involving the adequacy of the study blind and the other involving the fact that the parent ratings provided the only measure of behavior to assess treatment effects. With regard to blinding, while the investigators do state that the study was conducted under double-blind conditions (treatments provided to the parents in dated sealed envelopes) and that a special double capsule was used to administer treatment to the older ($\geq$6 years) children and orange juice with a straw was used as the vehicle for the younger children, there was no apparent effort to determine whether the capsules and orange juice adequately blinded the various dosages of tartrazine or to verify the effectiveness of the blinding procedures for the parents and children. However, with results showing clear dose response effects across a range of multiple dosage levels of tartrazine it seems likely that the integrity of the blind was adequately maintained in this study. With regard to the use of parental ratings as the sole outcome measure in this study, confidence in the reliability of study findings would be strengthened if other measures of behavior were also used to corroborate the parental findings. However, in view of the fact that the parent ratings were able to discern significant incremental differences in behavior and in clear dose related fashion across the dosage levels of tartrazine challenge, there appears to be sufficient confidence in the reliability of these study findings based solely on the parental behavioral ratings. Overall, these findings do indicate that tartrazine can cause in dose-related fashion across challenge doses of 1 to 50 mg adverse behavioral effects (consistently including irritability, restlessness and sleep disturbances and questionably associated with hyperactivity) in a subgroup of select children with problem behaviors who are suspected hyperactive and reportedly affected by diet.

(Schmidt et al, 1997)

Schmidt MH, Mocks P, Lay B, Eisert HG, Fojkar R, Fritz-Sigmund D, Marcus A and Musaeus B. Does oligoantigenic diet influence hyperactive/conduct-disordered children - a controlled trial. European Child & Adolescent Psychiatry, 1997, 6: 88-95.

A placebo controlled, double-blind dietary crossover study was conducted to examine the effectiveness of an oligoantigenic diet (eliminates all foods/additives assumed to provoke adverse behaviors in some children) in children diagnosed with ADHD/Disruptive Behavior Disorder, using multiple outcome measures of behavioral change. The dietary effects were also compared with the effects of stimulant medication treatment. Participants in this study consisted of an unselected sample of in-patients newly admitted to a hospital facility for treatment of hyperactive/disruptive behavior disorder. A total of 49 children (47 male, 2 female; 6 – 12 years) specifically diagnosed with ADHD and/or Conduct Disorder participated. The experimental diet was a very restrictive *oligoantigenic* diet that eliminates every kind of food and food additive (e.g. artificial colors) assumed to provoke behavioral reactions in certain children with beverages confined to apple juice and mineral water. The placebo control diet included common food items and ingredients, including food additives, assumed to be antigenic and also included daily beverages containing tartrazine (20 mg; FDC Yellow 5), quinoline yellow (10 mg), new coccine (10 mg) and carmoisine (10 mg). This placebo diet was specifically designed to control for the confounding effects of 'positive attention' in that the children were able to detect the difference in appearance between the experimental (oligoantigenic) and placebo diets, but were not supposed to be able to identify the experimental diet. It was not specified whether or not the children were actually able to identify the experimental diet. All personnel involved in rating behaviors were blind to diet treatments. All 49 children were tested with the experimental diet, but only 37 were additionally tested under non-blind conditions with stimulant medication for comparative purposes only. The stimulant medication was methylphenidate (0.4 mg /kg, orally); no placebo was specified since raters were not blind to medication treatment. Procedurally, each subject went through the following phases of test: *Phase1 /Pre-Study Baseline I* (4 days) - subjects were entered/acclimated to in-patient ward with standard food and assessment procedures (data not used for study); *Phase 2 / Diet Crossover Testing* (18 days @ 9 days per diet condition) - subjects were randomly assigned in pairs to the diet order of oligoantigenic/placebo or placebo/oligoantigenic and each diet was given for 9 days with two fixed assessment points at day 3 (T1) and day 8 (T2); *Phase 3 / Washout Baseline II* (9 days) – assessments were conducted once after at least 3 of the 9 washout days but diet condition (placebo control or regular clinic) was not specified; and *Phase 4 Methylphenidate Daily Treatment* ($\geq$ 5 days determined individually) - assessments were made in the morning of days 3-5 of drug treatment, 1.5 hours after drug ingestion, but diet condition was not specified. Outcome measures included both behavior ratings and objective laboratory testing. The behavior ratings were conducted in three standardized situations: in classroom, rated by teachers using the established widely used Conners Parent/Teacher 10 item rating scale; at play in the ward, rated by trained observers using a specially constructed 18 item behavior checklist; and during performance testing, rated by trained observers using a specially constructed 14 item behavior checklist. The laboratory testing consisted of two performance tests (continuous performance task or CPT, and paired associate learning task or PAT) and a continuous measure of activity (actometer). Blood levels of total IgE were determined once during baseline and once during the experimental diet testing. It should be noted that for purposes of data analysis the response definition (i.e., treatment effect) in this study was based only on the ratings data and did not include analysis of data from the objective laboratory testing. Since there were no significant

differences in behavior ratings between T1 and T2, these two assessment points were averaged for each child, and since there was no treatment order effect for the behavior ratings, these data for children of both crossover groups were pooled. Repeated measures analyses of variance were performed to analyze the data.

Study Results: With the response definition based only on the ratings data, group level analysis showed significant diet effects with improved behaviors rated by trained observers in both the play setting and the laboratory testing setting for those children on the oligoantigenic diet (9 days) compared with the placebo diet. However, there was no confirmation of a diet related effect on behavior based on the teacher ratings of classroom behavior. Objective laboratory testing results were not presented. Analysis of the rating data for individual subjects revealed significant behavioral improvement on the oligoantigenic diet for 22 of the 49 children (45%) in the laboratory testing setting and for 21 of the 49 children (43%) in the play setting. Parenthetically, significant behavioral deterioration on the oligoantigenic diet was also found for 7 (4%) children in the play setting and 6 (12%) in the laboratory testing setting. In terms of children exhibiting clinically significant changes, defined as 25% improvement on two measures, 12 of the 49 children (24%) were identified as "responders" to the oligoantigenic diet, and 2 (4%) showing more problem behaviors. The non-blind assessment of the effects of methylphenidate treatment ($\geq$ 5 days) was based on behavioral ratings in the play and laboratory testing settings only. Since the teacher ratings of classroom behavior did not reveal any diet effects, the investigators excluded classroom ratings from further group and individual analyses including those for the methylphenidate treatment. Therefore, the statement by the investigators that classroom ratings were not able to demonstrate the expected treatment effects of stimulant therapy is not completely accurate. Sixteen of 36 children (44%) with complete data showed clinically significant improvement in behavior (i.e., "responders") and 4 (11%) children showed deteriorated behavior with drug treatment. Overall, out of the 36 children who received both diet and methylphenidate testing, 11 (31%) improved on drug only, 3 (8%) improved on diet only, and 5 (14%) improved on both drug and diet. IgE levels, measured at baseline and under oligoantigenic diet condition, showed no differences between diet and baseline.

Study Assessment and Evaluation: In a sample of 49 children with ADHD/Conduct Disorder problem behaviors, 12 (24%) exhibited what was defined as clinically significant improvement in behavior when treated with an oligoantigenic diet for 9 days, based on behavior ratings by trained observers in two behavior settings (play and laboratory testing settings). But the teacher ratings in the classroom setting did not confirm any diet related effects on behavior. The investigators suggest that this negative finding in the classroom was the result of the teachers being highly experienced in dealing with psychiatric patients and that "biased ratings of a reduced range are, thus, rather to be expected". However, these results should be interpreted in light of the fact that the two behavioral rating checklists that were used in the play and laboratory testing settings were newly developed specifically for this study. While the investigators stated that on-site testing indicated both of these scales to be highly reliable and that significant correlation of both scales with activity measures appeared to support their validity, these newly devised rating scales have not been completely validated through use by other investigators. In contrast, the Conners behavioral scale used for the teacher rating in the classroom setting is a well-established, widely used, and validated rating scale. Although several objective laboratory tests were also conducted, unfortunately none of those data were presented or analyzed. Consequently, no additional information is available to corroborate any diet related effects. Regarding the relative effectiveness of diet and drug, although more patients appeared to respond to methylphenidate than to diet (44% and 24%, respectively), the level of improvement for those responders was comparable between both treatments. The absence of diet related differences in

IgE levels in this study do not support the notion that an IgE-related mechanism underlies the behavioral effects associated with food intolerance.

In addition to the inconsistent findings with significant diet related changes in behavior ratings by trained observers (play and laboratory testing settings) but no confirmation of diet effects in ratings by teachers (classroom setting), there are two other issues related to this study which significantly affect the confidence in the reliability of the study findings: (1) There is a possibility of a biasing influence that the clearly distinguished diets may have had on the behavioral ratings. No efforts were described to verify the effectiveness of the blinding, e.g. determine whether or not the children could actually tell which of the two diets was the experimental diet, whether they thought one of the diets should (adversely) affect their behavior, or whether the children told the raters which diet they were presently eating. (2) The response definition (i.e., treatment effect) in this study was based only on the behavioral ratings data and did not include data from the completed objective laboratory testing. The investigators considered the objective testing as less comprehensive, confounded by variables such as motivation, insensitive to subtle treatment effects, providing information already included in the behavioral ratings, not suitable for assessing response in a clinically relevant sense due to ceiling effects (CPT and PAT). It is unclear why these concerns were not considered relevant when these objective laboratory tests were included in and actually completed as part of the study design. The original inclusion of several objective laboratory tests in addition to the behavioral ratings was an experimental design feature that was viewed as adding credibility to the study by providing confirmatory evidence of treatment related effects. Simply dismissing these validly conducted test results *post hoc*, particularly without even presenting the data or their analyses, appears to be scientifically inappropriate and significantly lessens the confidence in the credibility of this study and in the completeness and reliability of the findings. Overall, considering the inconsistent behavioral findings (i.e., significant behavioral effects with ratings conducted in play and laboratory testing settings but no confirmation of behavioral effects with teacher ratings in the classroom), the use of incompletely validated behavioral checklists for ratings in play and laboratory test settings, the fact that objective laboratory tests (CPT, PALT and actometer readings) were collected but neither presented nor analyzed to help determine or confirm treatment related effects, together with the questionable effectiveness of the diet blinding conditions, these study results should be considered at best as only suggestive of a possible limited beneficial effect of the oligoantigenic diet in a small number of children with a profile of disruptive behaviors involving ADHD and Conduct Disorder. Alternatively expressed, this study may suggest that food intolerance may exaggerate some adverse behaviors in a small group of select children with ADHD/Conduct Disorder.

(Uhlig et al, 1997)

Uhlig T Merkenschlager A Brandmaier R Egger J. Topographic mapping of brain electrical activity in children with food-induced attention deficit hyperkinetic disorder. Eur J Pediatr, 1997 (July), 156(7):557-61.

A crossover clinical study was designed to investigate whether changes in brain electrical activity detectable by EEG topographical mapping occur in food sensitive ADHD children when exposed to provoking foods. The study was conducted in three phases - Phase I involved identifying ADHD children responsive to an oligoantigenic diet. The children participating in this phase were outpatients in a special dietary clinic and all were diagnosed ADHD with

Conners rating scores between 15 and 30 (6-15 years old; number and sex unspecified). For 3 weeks children received an open (non-blind) oligoantigenic (few foods) diet, which is a very restrictive diet that eliminates every kind of food and food additive assumed to provoke adverse behavioral reactions in certain children. Based on non-blind observations by parents and teachers and on Conners scores below 15 (the index level for ADHD), 71% of the children on diet showed improved behavior and were entered into the next phase of testing. Phase II involved the non-blind reintroduction of foods to identify specific provoking food items. With the children continuing on the oligoantigenic diet, eliminated foods were reintroduced one by one at the rate of one every 5 days. If hyperactive related behaviors did not occur that food item was incorporated into the diet. The provoking foods in order of prevalence were beet sugar, colors, wheat, cow milk, banana, egg, citrus, cacao, beef, pork, and oats. Phase III involved the experimental testing of provoking food items and EEG mapping. Fifteen children were entered into Phase III but only 12 completed testing. Procedurally, testing of each child was carried out both under "consuming" and "avoiding" provoking foods conditions. Eight children were tested in the order, "avoiding" then "consuming"; the other 4 children were tested in the reverse order, "consuming" then "avoiding". For the "avoiding" condition, children avoided provoking foods (oligoantigenic diet) for 14 days before EEG mapping. For the "consuming" condition, children ate normal helpings of their provoking foods daily (regular diet) for at least 5 days prior to and on the day of EEG mapping. Between the avoidance and consumption of provoking foods a 1-week washout period was interposed, although it was not stated what diet the children were receiving during washout. Outcome measures consisted of EEG recordings for topographic mapping conducted in 20 minute sessions at the end of each diet treatment condition and Conners behavior ratings conducted daily and during EEG recording sessions. Notably, only one study investigator was stated as being blind to order of testing but behavioral assessments were conducted by two investigators, and no other study personnel, including parents and children, were described as being blinded to treatment. So, it appears that the study was conducted virtually without effective blinding. The Wilcoxon two sample test was used to analyze data.

Study Results: Conners scores were significantly higher while the children were consuming the provoking foods. No order effect was detected. Significant differences between recordings after consumption or avoidance of the provoking food items were seen in the fronto-temporal regions of the brain with few changes in the parieto-occipital areas. Percentages of fast Beta-1 frequency electrical activity increased in frontal areas of the brain in association with abnormal behavior (elevated Conners scores) following consumption of provoking food items. The relative power spectra of all recordings while consuming provoking food items were significantly higher than during avoidance of the provoking food items.

Study Assessment and Evaluation: Unfortunately, in an otherwise well-conducted study, the critical absence of any apparent blinding conditions introduces the possibility of an unintentional influence on the behavior of the children, confounding interpretation of the EEG results and associated behavioral ratings. The children may have behaved differently not because of the diet but because they or their parents knew which diet they were receiving, which in turn could have affected the behavioral ratings and the EEG mapping. Also, the exclusion from the study of two subjects specifically because "there was no change in behaviour during the crossover study" is questionable. Overall, due primarily to the absence of blinding conditions, the results from this study cannot be interpreted to indicate any association between exposure to "provoking" foods (food intolerance) and specific EEG changes or behavioral rating changes. However, the results do appear to show credible evidence associating specific electrical activity changes in the frontal brain region with increased hyperactivity behaviors (indicated by increased Conners rating scores).

(Bateman et al, 2004)

Bateman B, Warner JO, Hutchinson E, Dean T, Rowlandson P, Gant C, Grundy J, Fitzgerald C and Stevenson J. The effects of a double blind, placebo controlled artificial food colourings and benzoate preservative challenge on hyperactivity in a general population sample of preschool children. Arch Dis Child, 2004 (June), 89(6): 506-11.

A double blind, placebo controlled challenge study was designed to determine whether food additives can cause hyperactive related behaviors in a sample of preschool children from the general population with and without hyperactivity and with and without atopy (exhibit allergic reactions). The 277 children who completed the challenge study were recruited from the general preschool population (Isle of Wright, UK) (average age 3 years) were assessed for hyperactivity and evaluated for atopy. ADHD was not clinically diagnosed but children were assessed for hyperactivity using two scales, completed by parents, which rated inattention, overactivity, fidgetiness, and impulsivity. Those children above a defined threshold scoring level were designated hyperactive (HA) and represented a distinct group of hyperactive 3 year old children. A wider range of behavioral problems were also assessed using the Behavioral Checklist scale. Children were defined as atopic (AT) if they showed an allergic reaction in a skin prick tests. The 277 children were sorted into 4 study groups: 36 HA/AT (22 male, 14 female), 75 HA/non-AT (39 male, 36 female), 79 non-HA/AT (44 male, 35 female), and 87 non-HA/non-AT (46 male, 41 female). The challenge item was a 300ml drink of mixed fruit juices with/without 20 mg mix of artificial food colors (5 mg each of sunset yellow, tartrazine, carmoisine, ponceau 4R) plus 45 mg sodium benzoate. The placebo and color drinks were in identical sealed bottles of the same appearance. Preliminary testing by 34 adults showed that the drinks could not be accurately differentiated on blind testing. All rating team members and family were blind to treatment. At the end of the study period the parents were equally divided into those who did or did not correctly identify the drink order (meaning that approximately half of the parents did correctly identify the drink order). Procedurally, after baseline data collection, all children were placed on a diet free of artificial color and benzoic acid for four weeks. Weeks 1 and 3 were "washout" weeks with no treatments. During weeks 2 and 4 all children received daily drinks of either placebo fruit juice or color challenge fruit juice (one type treatment per week determined randomly). Drinks were to be taken at home over the course of the day. Apparently, no drinks were given on the other study weeks. The two primary behavioral outcome measures included: (1) weekly clinical tests analyzed as an Aggregated Test Hyperactivity score (ATH), reflecting task performance and psychologist recording of behavioral measures of inattention, activity, and impulsivity; and (2) weekly parent ratings of behaviors from the Weiss-Werry-Peters activity scale analyzed as an Aggregated Parental Hyperactivity Rating score (APHR). In a preliminary analysis the clinical testing, including the ATH score, was validated as reliably distinguishing the cognitive and behavioral differences between hyperactive and non-hyperactive preschool children at baseline. Analysis of variance and analysis of covariance was used for to analyze the data.

Study Results: APHR (parents ratings) showed an overall significantly greater increase in hyperactivity for the active color challenge than for the placebo but there was no interaction with either initial hyperactivity or atopy (i.e., based on parent behavior ratings, adverse behavioral response to color/benzoate challenge occurred in both hyperactive and non-hyperactive children and in both allergic and non-allergic children). Curiously, there were increased levels of hyperactivity behavior for both the placebo and color/benzoate challenge periods, but a greater

increase in hyperactivity during the active periods. However, no behavioral effects of color/benzoate challenge were found with the ATH/clinical measures.

Study Assessment and Evaluation: In a group of 277 preschool children from the general population with/without hyperactivity and with/without atopy, overall group analysis of parent ratings indicated that the children exhibited a significantly greater increase in hyperactivity behaviors when challenged daily for one week with 20 mg artificial colors and 45 mg benzoic acid than with placebo treatment (both active and placebo treatments increased hyperactive behaviors, but active challenge a greater significant affect). This effect was seen for both the hyperactive and non-hyperactive children and for both the atopic and non-atopic children. However, it is not known what percent of the children were reported to be responsive to the color/benzoate challenge. In contrast to the parental ratings, the clinical testing measures, previously validated in-house for distinguishing between hyperactive and non-hyperactive children, provided no confirmatory evidence of any behavioral changes associated with the color/benzoate challenge in any of the children. In evaluating these contradictory findings, consideration should be given to the fact that, while efforts were made to maintain the blinding of all study participants, the interviews at the end of the study period found that the parents were equally divided into those who did or did not correctly identify the challenge treatment order. This would indicate that approximately half of the parents may not have been completely blind to treatment conditions. There would be more confidence in the parental ratings if they could be supported by observations by independent observers or by the use of standardized tests. In view of the equivocal finding of treatment effects between the two primary outcome measures, together with the uncertain effectiveness of the blind, there is limited confidence in the reliability of this study's findings. This study provides no support for the suggestion that food additives cause hyperactivity in preschool children from the general population. At best this study may suggest that non-allergy related intolerance to artificial food color additives and benzoic acid may occur in some children with or without hyperactivity related behaviors and that this intolerance may be associated with only limited behavioral changes.

(McCann et al, 2007)

McCann D, Barrett A, Cooper A, Crumpler D, Dalen L, Grimshaw K, Kitchin E, Lok K, Porteous L, Prince E, Sonuga-Barke E, Warner JO, Stevenson J. Food additives and hyperactive behaviour in 3-year-old and 8/9-year-old children in the community: a randomized, double-blinded, placebo-controlled trial. Lancet, 2007 (November 3), 370 (9598):1560-7.

McCann et al (2007) conducted a controlled challenge study to test whether intake of artificial food colors and additives (AFCA) affected childhood behaviors in a general population of preschool and school aged children. This study was designed in part to replicate a previous study on preschool children, to extend the investigation to school aged children, and to use a wider range of behavior measures. Two age groups of children, 3-4 year-old (preschool) and 8-9 year old (school aged), from the general population in Southampton, UK were recruited for participation. The initial samples consisted of 153 preschool children (79 male, 74 female) and 144 school aged children (75 male, 69 female). Neither group was diagnosed hyperactive or identified as having appreciable problem behaviors. The challenge material was a fruit juice drink containing one of two different color mixes plus sodium benzoate (Mix A and B). Two sets of each Mix were used, each set consisting of a Mix with lower color levels and a Mix with

higher color levels (lower levels for challenging preschool children and higher levels for challenging school aged children). Mix A/preschool contained 20 mg mixed color (5 mg sunset yellow; 2.5 mg carmoisine; 7.5 mg tartrazine; 5 mg ponceau 4R) and 45 mg sodium benzoate. Mix B/preschool contained 30 mg mixed color (7.5 sunset yellow; 7.5 mg carmoisine; 7.5 mg quinoline yellow; 7.5 mg allura red AC) and 45 mg sodium benzoate. Mix A/school-aged contained 25 mg mixed colors (levels of same colors increased by 1.25 times; sodium benzoate remained at 45 mg). Mix B/school-aged contained 62 mg mixed colors (levels of each color increased by 2.08 times; sodium benzoate remained at 45 mg). The preschoolers received 300 ml juice/day and the school aged children 625 ml/day. The placebo was mixed fruit juice. Families and all members of research team were blind to treatment. To test the challenge blind, two panels of 20 adult subjects were tested and reportedly were not able to distinguish the color mixes and placebo. Parents and children were not asked to verify blinding.

Procedurally, the study was conducted as a within-subject crossover challenge between two active color/benzoate mixes (A and B) and a placebo drink. During the initial baseline week (W0) regular diets were used, but during the subsequent 6 study weeks (W1-W6) a 'special diet' was used from which the additives used in the challenges (artificial colors and sodium benzoate) were withdrawn and the following weekly schedule of treatment was followed for all children: W1- washout week with daily placebo, W2 - daily active challenge or placebo, W3 - washout week with daily placebo, W4 - daily active challenge or placebo, W5 - washout week with daily placebo, W6 - daily active challenge or placebo. Challenge weeks were randomized for 2 active challenges (Mix A and Mix B) and 1 placebo treatment. Three sources of outcome behavioral measures were used for both the 3 year old and the 8/9 year old groups: (1) a weekly teacher ADHD rating, (2) a weekly parent hyperactivity rating (Weiss-Werry-Peters scale for preschool children and an ADHD rating scale IV for school aged children), and (3) a weekly trained observer classroom observation rating. One additional objective source measure, used for the school aged children only, was a weekly assessment of attention and response inhibition with the Conners continuous performance test. A single Global Hyperactivity Score (GHA) was calculated as a composite score of all test measures and used to assess individual differences in behavior based on the collective information from the different sources of outcome measures (teacher, parent, trained observers, and computerized laboratory test) covering various components of hyperactivity (overactivity, impulsivity, inattention). Because of differences between the testing of preschool and school aged children in the composition of the GHA scores and in the levels of colors used, the data for the two age groups could not be analyzed jointly. So, for purposes of data analysis the studies with the two age groups were treated as parallel but independent. Linear mixed-model methods were used to analyze the data and potential confounding factors were controlled. The data analyses were conducted for the full sample of children (the principle analysis for this study's findings) and replicated for a high consumption group (>85% consumption of drinks in any challenge week) and a complete case group (high consumption and no missing GHA scores). The latter two analyses were intended to determine whether non-compliance and the method of handling missing data affected the pattern of results.

Study Results: Based on differences in the composite GHA scores for the full sample, the young preschool children responded to Mix A challenge with significantly elevated hyperactivity scores but did not respond to Mix B challenge. Contrarily, the older school aged children responded to Mix B challenge with significantly elevated hyperactivity scores but did not respond to Mix A challenge. In terms of the magnitude of the behavioral changes, the effect sizes in this study averaged at about 0.18. The specific nature of the treatment related behavioral effects, generally referred to as increases in the mean level of hyperactivity, were not described. Also, other than the composite GHA scores, the investigators did not present the results for any of the component

outcome measures, i.e. teacher scores, parent scores, trained observer scores, and computerized laboratory testing.

Study Assessment and Evaluation: The fact that significant behavioral effects were found for the preschool children with Mix A but not Mix B and for the school aged children with Mix B but not Mix A seems to be a puzzling inconsistency. The investigators even acknowledge the need for additional study ”to establish whether the age-related differences seen in the present study can be replicated”. Also, the specific nature of the behavioral effects were not described, other than being generally referred to as increases in the mean level of hyperactivity. The children used in this study were from the general population and were not specifically diagnosed as ADHD. Consequently, the behaviors measured in this study, even though treatment effects were referred to by the investigators as increases in levels of "hyperactivity", were probably not the disruptive hyperactivity behaviors of ADHD but more likely the type of hyperactivity exhibited occasionally by the general population of preschool and school aged children. Furthermore, the magnitude of whatever behavioral changes may have been associated with the active challenges appeared to be rather small, with the effect sizes in this study averaging about 0.18 which is approximately equivalent to less than 10% of the difference between children with ADHD and without that disorder. Other than the composite GHA scores, the investigators did not present the results for any of the component outcome measures, i.e. teacher scores, parent scores, trained observer scores, and computerized laboratory testing, to enable assessing whether the various source measures were equally sensitive in detecting a treatment related change in behavior. In fact, information obtained by this reviewer from an earlier critique of the final report for this study showed that the component parent ratings appear to be one of the primary sources detecting treatment effects in both age groups of children but that the teacher and trained observer ratings were less able to detect significant treatment effects. It might have been more likely for teachers to have observed challenge related effects on behavior since many but not all children drank the juice challenge in the morning at breakfast and the children might have been expected to exhibit effect, in any occurred, during the school day. Unfortunately, this study appeared to lack control over when the challenges were ingested relative to timing of measures of behavior. The computerized testing in the older children did appear to detect some challenge related effects but there were unanticipated general performance decrements with repeated testing ostensibly attributed to boredom with that testing paradigm. The prominence of parental ratings, but not other clinical test measures, in detecting treatment related behavioral effects was also reported in the study by Bateman et al (1994). Overall, the primary study findings are suggestive of low level behavioral effects of a one week exposure to AFCAs on behavior in 3/4 year old children (Mix A) and 8/9 year old children (Mix B), based solely on parental ratings. However, due to the study weaknesses and caveats described above, particularly the puzzling inconsistency in response by the two age groups of children to both active challenges and the rather small treatment related changes, the reported effects appear to be equivocal and of uncertain biological relevance. This study does not provide credible evidence of deleterious effects of artificial food colors and additives on children’s behavior nor does it lend any support for the contention that food additives exacerbate hyperactive behaviors (inattention, impulsivity, overactivity) in children.

THE FINDINGS AND CONCLUSIONS IN THIS REPORT HAVE NOT BEEN FORMALLY DISSEMINATED BY FDA AND SHOULD NOT BE CONSTRUED TO REPRESENT ANY AGENCY DETERMINATION OR POLICY.

BIBLIOGRAPHY
All References Related to Overview Assessment of Artificial Food Colors/Additives and Hyperactivity (ADHD) and Problem Behaviors in Children

(I) Clinical Trials
- **(Ia) Reviewed Clinical Trials Prior to 1982 (n=16)**
- **(Ib) Reviewed Clinical Trials From 1982 to Present (n=17)**
- **(Ic) References Excluded from Overview (n=2)**

(II) Animal/Laboratory Studies and Background Information
- **(IIa) Animal/Laboratory Studies (n=17 referenced in review)**
- **(IIb) Background Information (n=51)**

I CLINICAL TRIALS

Ia. Reviewed Clinical Trials Published Prior To 1982 (n=16)

1. Adams W. Lack of behavioral effects from Feingold diet violations. Perceptual and Motor Skills, 1981, 52: 307-313.

2. Conners CK, Goyette CH and Newman EB. Dose-Time Effect of Artificial Colors. Journal of Learning Disabilities, 1980 (November),13 (No.9):48-52.

3. Conners CK, Goyette CH, Southwick DA, lees JM, *and* Andrulonis PA. Food additives and hyperkinesis: A controlled double-blind experiment. Pediatrics, 58 (August, No.2), 1976: 154-166.

4. Conners CK. Artificial colors in the diet and disruptive behavior. In R Knights and D Bakker (Eds.) *Treatment of Hyperactive and Learning Disordered Children: Current Research*. Baltimore, University Park Press, 1980: 113-120.

5. Goyette CH, Connors CK, Petti TA and Curtis LE. Effects of Artificial Colors on Hyperkinetic Children: A Double-Blind Challenge Study. Psychopharmacology Bulletin, 14 (No.2), 1978: 39-40.

6. Harley JP, Ray RS, Tomasi L, Eichman PL, Matthews CG, Chun R, Cleeland CS and Traisman E. Hyperkinesis and food additives: testing the Feingold hypothesis. Pediatrics. 1978, Jun; 61(6):818-28

7. Harley P, Matthews CG and Eichman P. Synthetic Food Colors and Hyperactivity in Children: Double-Blind Challenge Experiment. Pediatrics, 62, 1978: 975-983.

8. Levy F and Hobbes G. Hyperkinesis and Diet: A Replication Study in Hyperactive Children. American Journal of Psychiatry. 1978:135: 1559-1560.

9. Levy F, Dumbrell S, Hobbes G, Ryan M, Wilton N and Woodhill JM. Hyperkinesis and Diet: A Double-Blind Crossover Trial With A Tartrazine Challenge. Medical Journal of Australia, 1 (No.2, January), 1978: 61-64.

10. Mattes J and Gittelman R. Effects of Artificial Food Colorings In Chldren With Hyperactive Symptoms: A Critical Review and Results of a Controlled Study. Archives of General Psychiatry, 38, 1981:714-718.

11. Mattes J and Gittelman-Klein R. A Crossover Study of Artificial Food Colorings in a Hyperkinetic Child. American Journal of Psychiatry, 135 (No.8), 1978: 987-988.

12. Rose TL. The Functional Relationship Between Artificial Food Colors and Hyperactivity. Journal of Applied Behavior Analysis, 115 (No.4, Winter), 1978: 439-446.

13. Spring C, Vermeersch J, Blunden D and Sterling H. Case Studies of Effects of Artificial Food Colors on Hyperactivity. The Journal of Special Education, 1981, 15 (Number 3): 361-372.

14. Swanson JM and Kinsbourne M. Food dyes impair performance of hyperactive children on a laboratory learning test. Science, 1980, 207 (March 28): 1485-1487.

15. Weiss B, Hicks William J, Margen S, Abrams B, Caan B, Citron LJ, Cox C, McKibben J, Ogan D and Schultz S. Behavioral Responses to Artificial Food Colors. Science, 207 (March), 1980: 1487-1489.

16. Williams JI, Cram DM, Tausig FT and Webster E. Relative Effects of Drugs and Diet on Hyperactive Behaviors: An Experimental Study. Pediatrics, 61, 1978: 811-817.

Ib. Reviewed Clinical Trials Published From 1982 to Present (n=17)

17. Bateman B, Warner JO, Hutchinson E, Dean T, Rowlandson P, Gant C, Grundy J, Fitzgerald C and Stevenson J. The effects of a double blind, placebo controlled, artificial food colourings and benzoate preservative challenge on hyperactivity in a general population sample of preschool children. Arch Dis Child. 2004, Jun; 89(6):506-11.

18. Boris M and Mandel FS. Foods and additives are common causes of the attention deficit hyperactive disorder in children. Ann Allergy. 1994, May; 72(5):462-8.

19. Carter CM, Urbanowicz M, Hemsley R, Mantilla L, Strobel S, Graham PJ, Taylor E. Effects of a few food diet in attention deficit disorder. Arch Dis Child. 1993, Nov; 69(5):564-8.

20. David TJ. Reactions to dietary tartrazine. Archives of Disease in Childhood,62,1987:119-122.

21. Egger J, Carter CM, Graham PJ, Gumley D andSoothill JF. Controlled trial of oligoantigenic treatment in the hyperkinetic syndrome. Lancet. 1985, Mar 9; 1(8428):540-5

22. Egger J., Stolla A and McEwen LM. Controlled trial of hyposensitization in children with food-induced hyperkinetic syndrome. The Lancet, 339 (May), 1992: 1150-1153.

23. Kaplan BJ, McNicol J, Conte RA and Moghadam HK. Dietary replacement in preschool-aged hyperactive boys. Pediatrics, 83 (No.1, January), 1989: 7-17.

24. McCann D, Barrett A, Cooper A, Crumpler D, Dalen L, Grimshaw K, Kitchin E, Lok K, Porteous L, Prince E, Sonuga-Barke E, Warner JO, Stevenson J. Food additives and hyperactive behaviour in 3-year-old and 8/9-year-old children in the community: a randomised, double-blinded, placebo-controlled trial. Lancet. 2007, Nov 3; 370(9598):1560-7.

25. Pollock I, Warner JO. Effects of artificial food colours on childhood behaviour. Arch Dis Child. 1990, Jan; 65(1):74-7.

26. Rowe KS, Rowe KJ. Synthetic food coloring and behavior: a dose response effect in a double-blind, placebo-controlled, repeated-measures study. J Pediatr. 1994, Nov; 125(5 Pt 1):691-8

27. Rowe KS. Synthetic food colourings and 'hyperactivity': a double-blind crossover study. Aust Paediatr J. 1988, Apr; 24(2):143-7.

28. Salamy J,Shucard D,Alexander H, Peterson D, Braud L. Physiological changes in hyperactive children following the ingestion of food additives. Int J Neurosci. 1982, May; 16(3-4):241-6.

29. Sarantinos J, Rowe KS and Briggs DR. Synthetic Food Colouring and Behavioral Change In Children With Attention Deficit Disorder: A Double-Blind, Placebo Controlled (Challenge Study). Proc. Nutr. Aust., 15, 1990: 233.

30. Schmidt MH, Mocks P, Lay B, Eisert HG, Fojkar R, Fritz-Sigmund D, Marcus A and Musaeus B. Does oligoantigenic diet influence hyperactive/conduct-disordered children - a controlled trial. European Child & Adolescent Psychiatry, 6, 1997: 88-95.

31. Thorley G. Pilot study to assess behavioural and cognitive effects of artificial food colours in a group of retarded children. Dev Med Child Neurol. 1984, Feb; 26(1):56-61.

32. Uhlig T Merkenschlager A Brandmaier R Egger J. Topographic mapping of brain electrical activity in children with food-induced attention deficit hyperkinetic disorder. Eur J Pediatr. 1997, Jul; 156(7):557-61.

33. Wilson N and Scott A. A double-blind assessment of additive intolerance in children using a 12 day challenge period at home. Clinical and Experimental Allergy, 19, 1989: 267-272.

Ic. References Originally Listed but Excluded from Overview (n=2)

34. Harley JP, Tomasi L, Ray R, Eichman P, Matthews C, Chun R, Traisman E and Cleeland C. An experimental evaluation of hyperactivity and food additives – Phase I. Univ Wisconsin Madison Rep, 207, 1977. *REASON NOT USED* : this reference was included in the original list of clinical study references but was excluded from the overview since it was only an early outline of the fully published study (Harley et al, 1978) which was reviewed.

35. Silfverdal SA, Hernell O. Food additives can increase hyperactivity in children. Results from a British study confirm the connection. Lakartidningen. 2008 Feb 6-12; 105(6):354-5. *REASON NOT USED* : this is a commentary and not a clinical study.

II ANIMAL LABORATORY STUDIES and BACKGROUND INFORMATION

IIa. Animal Laboratory Studies (n=17 referenced in review)

36. Chambille I. Retinal ganglion cells expressing the FOS protein after light stimulation in the Syrian hamster are relatively insensitive to neonatal treatment with monosodium glutamate. J Comp Neurol. 1998, Mar 23; 392(4):458-67. (Study not relevant to this overview – study does not include food colors and dose of MSG is clearly neurotoxic/neuropathologic at 400 g/kg)

37. Clough RW, Aravich PF, Sladek CD. Monosodium glutamate neurotoxicity a sex-specific impairment of blood pressure but not vaspressin in developing rats. Brain Res Bull, 17 (No.1), 1986 : 51-58. (Study not relevant to this overview - study does not include food colors and dose of MSG is clearly neurotoxic/neuropathologic at 400 g/kg;also, the study findings are questionable due to the very small n and the fact that all rats apparently came from the same few litters such that litter effects were not controlled.)

38. Dawson, R Jr, Simkins JW and Wallace DR. Age and dose-dependent effects of neonatal monosodium glutmate MSG administration to female rats. Neurotoxicol Teratol, 11(No.4), 1989 : 331-338. (Study not relevant to this overview – study does not include food colors and dose of MSG is clearly neurotoxic/neuropathologic at 400 g/kg)

39. Ford GP, Stevenson BI and Evans JG. Long-term toxicity study of carmoisine in rats using animals exposed in utero. Food Chem Toxicol. 1987, Dec; 25(12):919-25.

40. Galloway WD, Olvey KM and Brown NT. Behavioral effects of erythrosine following light exposure. Neurobehav Toxicol Teratol. 1986 Sep-Oct; 8(5):493-7.

41. Goldenring JR, Batter DK and Shaywitz BA Effect of chronic Erythrosin B administration on developing rats. Neurobehav Toxicol Teratol. 1981 Jan-Feb; 3: 57-58.

42. Goldenring JR, Batter DK and Shaywitz BA. Sulfanilic acid: behavioral change related to azo food dyes in developing rats. Neurobehav Toxicol Teratol. 1982 Jan-Feb; 4(1):43-9.

43. Kiss A, Skultetyova I, Jezova D. Corticotropin-releasing hormone synthesizing neurons in the hypothalamic paraventricular nucleus of rats neonatally treated with monosodium glutamate can respond to different stress paradigms. Neurol Res. 1999, Dec; 21(8):775-80. (Study not relevant to this overview – study does not include food colors and dose of MSG is clearly neurotoxic/neuropathologic at 400 g/kg)

44. Kiss P Hauser D Tamas A Lubics A Racz B Horvath ZS Farkas J Zimmermann F Stepien A Lengvari I Reglodi D. Changes in open-field activity and novelty-seeking behavior in periadolescent rats neonatally treated with monosodium glutamate. Neurotox Res. 2007, Sep; 12(2):85-93. (Study not relevant to this overview – study does not include food colors and dose of MSG is clearly neurotoxic/neuropathologic at 400 g/kg)

45. Kluwe WM Montgomery CA Giles HD Prejean JD. Encephalopathy In Rats And Nephropathy In Rats And Mice After Subchronic Oral Exposure To Benzaldehyde. Food and Chemical Toxicology, 1983, 21(No. 3): 245-250.

46. Lafferman JA and Silbergeld EK. Erythrosine B inhibits dopamine transport in rat caudate synaptosomes. Science, 1979 (July), 205: 410-412. (Reviewed in MacGibbon, 1983; Mailman and Lewis, 1983; Silbergeld and Anderson, 1981.)

47. Lau K McLean WG Williams DP Howard CV. Synergistic interactions between commonly used food additives in a developmental neurotoxicity test. Toxicol Sci. 2006, Mar; 90(1):178-87.

48. Levitan H, Ziylan YZ and Rapoport SI. Brain uptake of the food dye erythrosine B. Int Res Commun Syst Med Sci, 13 (No.1), 1985: 64.65.

49. Maher TJ, Wurtman RJ. Possible neurologic effects of aspartame, a widely used food additive. Environ Health Perspect. 1987, Nov; 75:53-7. (Study not relevant to this overview – study does not include food colors)

50. Mailman, RB, Ferris RN, Tang GLM, et al. Erythrosine (Red No. 3) and its nonspecific biochemical actions: what relation to behavioral changes? Science, 1980, 207 : 535-557.

51. Sobotka TJ, Brodie RE and Spaid SL. Tartrazine and the developing nervous system of rats. J Toxicol Environ Hlth, 1977, 2 : 1211-1220.

52. Sun YM, Ho ML, Hsu HK and Peng MT. The effects of neonatal monosodium glutamate treatment on sex-odor attractivity and approach behavior in rats. KAOHSIUNG J MED SCI; 9 (4). 1993. 232-242 (Study not relevant to this overview – study does not include food colors and dose of MSG is clearly neurotoxic/neuropathologic at 400 g/kg)

53. Tanaka T. Effects of amaranth on F1 generation mice. Toxicol Lett. 1992, May; 60(3):315-24.

54. Tanaka T. Reproductive and neurobehavioral effects of Allura Red AC administered to mice in the diet. Toxicology. 1994, Sep 6; 92(1-3):169-77.

55. Tanaka T. Reproductive and neurobehavioural effects of lac dye administered in the diet to mice. Food Addit Contam. 1997 May-Jun; 14(4):373-80.

56. Tanaka T. Reproductive and neurobehavioural toxicity study of erythrosine administered to mice in the diet. Food Chem Toxicol. 2001, May; 39(5):447-54.

57. Tanaka T. Reproductive and neurobehavioural toxicity study of tartrazine administered to mice in the diet. Food Chem Toxicol. 2006, Feb; 44(2):179-87.

58. Vorhees CV Butcher RE Brunner RL Wootten V Sobotka TJ. Developmental toxicity and psychotoxicity of FD and C red dye No. 40(allura red AC) in rats. Toxicology. 1983; 28(3):207-17.

59. Vorhees CV Butcher RE Brunner RL Wootten V Sobotka TJ. A developmental toxicity and psychotoxicity evaluation of FD and C red dye #3 (erythrosine) in rats. Arch Toxicol. 1983, Aug; 53(4):253-64.

IIb. Background Information (n=51)

60. Arnold LE Treatment alternatives for Attention-Deficit Hyperactivity Disorder (ADHD). Journal of Attention Disorders, 3 (No.1/April). 1999: 30-48.

61. Banerjee TD Middleton F Faraone SV Environmental risk factors for attention-deficit hyperactivity disorder. Acta Paediatr. 2007, Sep; 96(9):1269-74.

62. Bishop CL Attention deficit disorder and the Feingold diet Can. J. Hosp. Pharm.; VOL 36 ISS Fall 1983, P71-74(REF 17).

63. Brenner A Food additives and behavior. Md Med J.1986,May; 35(5):344-5.

64. Brookes KJ, Mill J, Guindalini C, Curran S, Xu XH, Knight J, et al. A common haplotype of the dopamine transporter gene associated with attention-deficit/hyperactivity disorder and interacting with maternal use of alcohol. Archives of General Psychiatry. 2006, 63: 74-81. (Reviewed in Banerjee at al, 2007; Stevenson, 2006.)

65. Center for Science in the Public Interest. Diet, ADHD & Behavior/A Quarter-Century Review: 2008 Update on Food Dyes and Behavior. Center for Science in the Public Interest. Washington, DC 20009.

66. Conners CK (b). Symptom patterns in hyperkinetic, neurotic, and normal children. Child Development, 1970, 41:667-682.

67. Cormier E Elder JH. Diet and child behavior problems: fact or fiction?Pediatr Nurs. 2007 Mar-Apr; 33(2):138-43

68. Cruz NV Bahna SL Do food or additives cause behavior disorders? Pediatr Ann. 2006, Oct; 35(10):744-5, 748-54.

69. Edwards AM. Food-allergic disease. Clinical and Experimental Allergy, 1995, 25 (Suppl. 1): 16-19.

70. Faraone SB, Dolye AE, Mick E and Biederman J. Meta-analysis of the association between the 7 repeat allele of the dopamine d4 receptor gene and Attention-Deficit Hyperactivity Disorder. American Journal of Psychiatry. 2001, 158 : 1052-1057. (Reviewed in Stevenson, 2006)

71. Faraone SV, Perlis RH, Doyle AE, Smoller JW, Goralnick J, Holmgren MA, et al. Molecular genetics of attention deficit hyperactivity disorder. Biol Psychiatry. 2005 57:1313-1323. (Reviewed in Banerjee et al, 2007)

72. Feingold B. *Hearings on Food Additives and Hyperactivity in Children*. Congressional Record, U.S. Senate, October 30, 1973. (Reviewed in Schauss, 1983)

73. Feingold, B. 1975 Hyperkinesis and rearing disabilities linked to artificial food flavors and colors. Am J Nurs, 1975, 75:797. (Reviewed in Mattes, 1983; Thorley, 1983; Stevenson, 2006)

74. Goodman R and Stevenson J. A twin study of hyperactivity number 2: The aetiological role of genes, family factors and perinatal adversity. Journal of Child Psychology and Psychiatry. 1989, 30: 691-709. (Reviewed in Stevenson, 2006)

75. Kavale KA, Forness SR. Hyperactivity and diet treatment: a meta-analysis of the Feingold hypothesis. J Learn Disabil. 1983 Jun-Jul; 16(6):324-30.

76. Kidd PM Autism, an extreme challenge to integrative medicine. Part 2: medical management. Altern Med Rev. 2002, Dec; 7(6):472-99.

77. King DS Psychological and behavioral effects of food and chemical exposure in sensitive individuals. Nutr Health. 1984; 3(3):137-51.

78. Krummel DA Seligson FH Guthrie HA Hyperactivity: is candy causal? Crit Rev Food Sci Nutr. 1996, Jan; 36(1-2):31-47.

79. Lipton MA Mayo JP Diet and hyperkinesis--an update. J Am Diet Assoc. 1983, Aug; 83(2):132-4.

80. Lucarelli S, Frediani T, Zingoni AM, Ferruzzi F, Giardino O, Qunitieri F, Barbato M, D'Eufemia P, and Cardi E. Food allergy and infantile autism.PANMINERVA MEDICA; 37 (3). 1995. 137-141.

81. MacGibbon B Adverse reactions to food additives. Proc Nutr Soc. 1983, Jun; 42(2): 233-240.

82. Mailman RB Lewis MH Food additives and childhood hyperactivity. ASDC J Dent Child. 1983 Jul-Aug; 50(4):283-6.

83. Mattes JA The Feingold diet: a current reappraisal. J Learn Disabil. 1983 Jun-Jul; 16(6):319-23.

84. National Institutes of Health. Diagnosis and Treatment of Attention Deficit Hyperactivity Disorder: NIH Consensus Statement. 1998 (November 16-18), 16: 1-37.

85. National Institutes of Health. Defined Diets and Childhood Hyperactivity: NIH Consensus Statement. (1)Am J Clin Nutr. 1983 (Jan), 37(1):161-5.; and (2)Clin Pediatr (Phila). 1982 (Oct 21), 10:6 27-30.

86. Pollock I Hyperactivity and food additives. Bibl Nutr Dieta. 1991(48):81-9.

87. Rapp DJ. Food additives and hyperactivity : Letter. LANCET 1982, 1(8281) 1128

88. Ribon A, Joshi S Is there any relationship between food additives and hyperkinesis? Ann Allergy. 1982, May; 48(5):275-8.

89. Rimland B The Feingold diet: an assessment of the reviews by Mattes, by Kavale and Forness and others. J Learn Disabil. 1983 Jun-Jul; 16(6):331-3

90. Rippere V Food additives and hyperactive children: a critique of Conners. Br J Clin Psychol. 1983, Feb; 22 Pt 1:19-32.

91. Robinson J and Ferguson A. Food sensitivity and the nervous system. Nutr Res Rev, 1992, 5: 203. (Reviewed in Young, 1997)

92. Schauss AG. Nutrition and behavior: Complex interdisciplinary research. Nutrition and Health. 1984, 3: 9-37.

93. Schab DW Trinh NT Do artificial food colors promote hyperactivity in children with hyperactive syndromes? A meta-analysis of double-blind placebo-controlled trials. J Dev Behav Pediatr. 2004 (December), 25(6):423-34.

94. Schnoll R, Burshteyn D, Cea-Aravena J. Nutrition in the treatment of attention-deficit hyperactivity disorder: a neglected but important aspect. Appl Psychophysiol Biofeedback. 2003, Mar; 28(1):63-75.

95. Shaywitz BA Sullivan CM Anderson GM Gillespie SM Sullivan B Shaywitz SE Aspartame does not affect behavior and cognitive function in children with Attention-Deficit Disorder (ADD). 103rd Annual Meeting of the American Pediatric Society and 62nd Annual Meeting of the American Society for Pediatric Research. Washington, D.C. May 3-6, 1993. Pediatric Research, 1993, 33 (4 Part 2): 17A.

96. Silbergeld EK Anderson SM Artificial food colors and childhood behavior disorders.Bull N Y Acad Med. 1982, Apr; 58(3):275-95.

97. Sonuga-Barke EJS. The dual pathway model of AD/HD: an elaboration fo neuro-developmental characteristics. Neuroscience and Biobehavioral Reviews. 2003, 27: 593-604. (Reviewed in Stevenson, 2006)

98. Stare FJ, Whelan EM, Sheridan M. Diet and hyperactivity: is there a relationship? Pediatrics, 1980, 66:521. (Reviewed in Bishop, 1983)

99. Stevenson J Dietary influences on cognitive development and behaviour in children. Proc Nutr Soc. 2006, Nov; 65(4):361-5.

100. Stevenson J, Sonuga-Barke E and Warner J. Chronic and acute effects of artificial colourings and preservatives on children's behaviour. Study Report: School of Psychology, University of Southampton (England). 2007 (b)

101. Story M Neumark-Sztainer D Diet and adolescent behavior: is there a relationship? Adolesc Med. 1998, Jun; 9(2):283-98, vi.

102. Taylor EA Attention deficit disorder and hyperkinesis.Indian J Pediatr. 1984 Mar-Apr; 51(409):193-204.

103. Thorley G Childhood hyperactivity and food additives. Dev Med Child Neurol. 1983, Aug; 25(4):531-4.

104. Watson R European agency rejects links between hyperactivity and food additives. BMJ. 2008, Mar 29; 336(7646):687

105. Weber RW Food additives and allergy. Ann Allergy. 1993, Mar; 70(3):183-90.

106. Weiss B Food additives as a source of behavioral disturbances in children. Neurotoxicology. 1986 (Summer); 7(2):197-208.

107. Wender EH The food additive-free diet in the treatment of behavior disorders: a review. J Dev Behav Pediatr. 1986, Feb; 7(1):35-42.

108. Wolraich ML, Lingren SD, Stumbo PJ, Stegink PJ, Appelbaum MI and Kiritsy MC. Effects of diets high in sucrose or aspartame on the behavior and cognitive performance of children. New England Journal of Medicine, 330 (No. 5), 1994: 301-307.

109. Wolraich ML, Wilson DB, White W. The effect of sugar on the beahvior or cognition of children: A meta-analysis. JAMA. 1995, 274: 1617.

110. Young E Prevalence of intolerance to food additives. Environmental Toxicology and Pharmacology, 1997, 4 (1-2). 111-114.

THE FINDINGS AND CONCLUSIONS IN THIS REPORT HAVE NOT BEEN FORMALLY DISSEMINATED BY FDA AND SHOULD NOT BE CONSTRUED TO REPRESENT ANY AGENCY DETERMINATION OR POLICY.

Table 1. Summary of Reviewed Clinical Trials Related to Artificial Food Colors and Problem Behaviors/ADHD in Children (Chronological Order)

Trials (Chron. Order)	N	Study Selection Criteria*			Treatment	Treatment Duration	Behavior Outcome Measures	Design Weaknesses/ Mitigating Factors	GROUP EFFECTS GROUP Analysis Treatment Effects for Various Outcome Measures						INDIVIDUAL Analysis Treatment Effects
		Behavioral Criteria	Age	Diet** Responsive					*Parent Rating*	*Teacher Rating*	*Classroom Assessm/ Other Observer Rating*	*Neuro-psych Testing*	*Activity*	*Other*	
Conners et al, 1976	15	Hyperactive ***	6-12 yrs	NA	F/KP diet	4-wk diet treatment	Conners/ Behavior Rating (Parent, Teacher)	*- treatment x order; unclear blinding effectiveness*	no diet effect	diet effect					4-5/15 affected w both parent & teacher ratings
Harley et al, 1978 (Phase I) Trial 1	36	Hyperactive	6-13 yrs	NA	F/KP diet	3-4 wks diet treatment	C/Behavior Rating (Parent, Teacher), Classroom Observations, and Neuropsychology Testing including activity	*- treatment x order; unclear blinding effectiveness*	diet effect	no diet effect	no diet effect	no diet effect	no diet effect		4/36 affected w parent & teacher ratings only
Harley et al, 1978 (Phase I) Trial 2	10	Hyperactive	3-6 yrs	NA	F/KP diet	3-4 wks diet treatment	C/Behavior Rating (Parent) and Neuropsychology Testing including activity measures	*- unclear blinding effectiveness*	diet effect			no diet effect	no diet effect		
Harley et al, 1978 (Phase II)	9	Hyperactive	10 yrs (from Phase I)	yes (F/KP)	candybars & cookies w ~13.5 mg color mix (2 items per challenge)	2 wks & 3 wks daily challenge	C/Behavior Rating (Parent, Teacher), Classroom Rating and Neuropsychological Tests		no color effect	no color effect	no color effect	no color effect			1/9 general pattern of effect w parent and classroom observer rating only
Goyette et al, 1978 Segment 1	16	Hyperactive	8 yrs (5-12)	yes (F/KP)	choc cookie w 26 mg color mix	2 x 2-wks daily challenge	C/Behavior Rating (Parent, Teacher) and Tracking Task	*- unclear blinding effectiveness*	no color effect	no color effect				a non-significant color effect trend for deficit in tracking test	

Goyette et al, 1978 Segment 2	13	Heterogeneous (8 Hyperactive / 5 Borderline Hyperactive)	3-10 yrs	yes / borderline (F/KP)	choc cookie w 13 mg color mix	1 wk daily challenge	C/Behavior Rating (Parents only ; Teacher ratings incomplete)	*- unclear blinding effectiveness; single outcome measure (parent rating).*	color mix challenge effect						
Levy and Hobbes, 1978	7	Hyperactive	5 yrs (6m,1f)	yes (F/KP)	cookie w 1 mg Tartrazine (Yellow 5) (4 cookies challenge)	2 wks daily challenge	C/Behavior Rating (Parent) and Parent Global Assessment	*- unclear blinding effectiveness; single outcome measure (parent rating)*	no tartrazine effect						
Levy et al, 1978	20	Hyperactive	4-8 yrs (19m,3f)	yes (F/KP), determined in open trial	biscuits @ 1 mg tartrazine (Yellow #5) (5 biscuits challenge)	2 wks daily challenge	C/Behavior Rating (Parent, Teacher, Clinician) and Multiple Psychometric Tests	*- blinding compromised; restricted subgroup analysis to single outcome measure (maternal ratings)*	no tartrazine effect	no tartrazine effect		no tartrazine effect	no tartrazine effect		subgroup rated by mothers as strong diet responders in open trial showed tartrazine challenge effects using maternal behavior rating only
Mattes and Gittelman-Klein, 1978	1	Hyperkinetic (primarily restless and irritable)	10 yrs	yes (F/KP)	Cookie w color mix at 1/5 ADI (3 cookies challenge based on dose-ranging)	2 days/wk on 5 alternating wks	C/Behavior Rating (Parent, Teacher)	*- single subject*	no color mix effect	no color mix effect					
Rose, 1978	2	Hyperactive	8 yrs (2f)	yes (F/KP)	oatmeal cookie w 1.2 mg Tartrazine (Yellow 5)	1 day/wk on 2 alternating wks	C/Behavior Assessment (Trained Observer of hyperactive behaviors in school setting)	*- only 2 subjects; unclear blinding effectiveness; single outcome measure (trained observer)*			tartrazine challenge effect				
Williams et al, 1978	26	Heterogeneous (19 hyperactive ; 7 non-HA w problem behaviors)	6-14 yrs	not diet responsive (kept on modified Feingold diet during test)	chocolate cookie w ~ 13 mg color mix (2 cookies challenge); stimulant drug	4 challenge days/wk over 2 wks (color+drug & color+vehicle)	C/Behavior Rating (Parent, Teacher)	*- inconsistent teacher vs parent ratings; treatment limited to 4 challenge days; marked variations in rating scores*	no color mix effect	color mix challenge effect					no agreement between teachers list of 5/26 responders and parents list of 3/26 responders

Conners, 1980 (replicate of Goyette et al, 1978 (trial II))	30	Heterogeneous	6-12 yrs	yes (F/KP) (determined 25% improved behavior in open diet trial)	Cchocolate cookie w 13 mg color mix (2 cookies per challenge)	daily challenge/wk on 2 alternate wks	C/Behavior Rating (Parent, Teacher)	*- unclear blinding effectiveness; single outcome measure (parent rating)*	no color mix effect	no color mix effect					
Conners et al, 1980 (time response)	9	Hyperactive	5-11 yrs (5m/4f)	yes (F/KP) (selected for marked behavioral changes with Feingold diet and notable sensitivity to artificial colors)	chocolate cookies w 15 mg color mix (2 cookies per challenge)	one-time color challenge	Behavior Rating (Experimenter) and Psychometric Testing (actometer, activity chair, paired associate attention & learning task)	*- confounding practice effect with learning task; low dose 30 mg challenge*			no color mix effect	no color mix effect	no color mix effect		
Swanson and Kinsbourne, 1980	40	Heterogeneous (20 hyperactive, drug responsive; 20 non-HA w problem behavior and adverse drug response)	10 yrs	not diet responsive	Capsules w color mix @ 100, 150 mg (placebo capsules w sugar)	one-time color challenge	C/Behavior Rating (Investigator) and Paired Associate Task (3 time intervals post challenge)	*- unclear blinding effectiveness; inconsistent placebo learning performance*			no color mix effect	color mix challenge effect for hyperactive but not non-hyperactive subjects			
Weiss et al, 1980	22	Non-Hyperactive w problem behaviors	3-7 yrs (15m,7f)	yes (F/KP)	Drink w 35.3 mg color mix	8 random days of challenge during 11 wk study	C/Behavior Rating (Parent), Individualized Behavior Rating (10 item) (Parent), Global Behavior Estimates (Parent), Sleep Record (Parent), Aversive Behavior Count (Parent)	*- unclear blinding effectiveness; single source outcome measure (parent ratings)*							Parental behavior ratings/assessments identified 1 out of 22 children with convincing adverse behavioral response to the color mix challenge
Mattes and Gittelman-Klein, 1981	11	Heterogenous (6 ADHD/ADD; 5 problem behaviors)	4-13 yrs	yes (F/KP)	Cookies w 13 mg color mix (incremental daily dosing w max 6 cookies per challenge @ 78 mg)	1- 6 cookies/day for 1 wk	Conners and Hyperactivity Behavior Ratings (Parents, Teacher), Psychiatric Evaluation, Distractability Test and Behavioral Rating (Psychologist)	*NONE*	no color mix effect	no color mix effect	no color mix effect	no color mix effect			Parents ratings identified 6/11 children as responding to the color mix challenge, but not confirmed by other outcome measures

Adams, 1981	18	Hyperactive	7.6 yrs (4-12) (15m/3f)	yes (F/KP) *(selected for diet response and negative effects with diet infractions)*	Chocolate cupcake & lemonade @ 26.3 mg mix of red/yellow food colors plus vanilla flavor and other unspecified additives	1 "artificial snack" 3-4 hours prior to testing	Battery of Objective Tests (activity, attention, motor skills) and Parents Rating of Degree of Behavioral Change	*- unclear blinding effectiveness*	no color/ additive "snack" effect			no color/ additive "snack" effects			
Spring et al, 1981	6	Hyperactive (confirmed w mothers' ratings)	8-13 yrs (m)	yes (F/KP) *(selected for mothers' claim of response to diet)*	Cookies w 13 mg color mix (two cookies per challenge = total daily challenge 26 mg)	3 daily challenges during each of 2 separate weeks	Modified Conners Hyperactivity Rating (Mothers/Teachers) ; Behavior-Based Guess of Which Treatment (Mothers/Teachers) [Replication Trial: Mothers' measures only]	*NONE*	inconsistent color mix effect	inconsistent color mix effect					
Salamy et al, 1982	8	Heterogeneous (4 hyperactive ; 4 non-HA siblings)	5-11 yrs	yes (F/KP)	Lemon drink w Red 40/artificial additives (level unspecified)	single one day challenge	Physiological Measures (EEG, HR)	*- inadequate test sessions; small N to assess variability; levels of test articles not specified*						no Red 40/additive effect on EEG or HR	
Thorley, 1984	10	Non-Hyperactive Retarded (w inattentive/ hyperactive behaviors)	10+ yrs	not diet responsive	Cocao drink w 92 mg 16-color mix	2 consecutive days challenge	C/Behavior Rating (Health Care Staff, Teacher), Individualized/ Behavior Rating (Teacher), Psychometric Testing (mazes, paired associates test, actometer)		no color mix effect	no color mix effect		no color mix effect	no color mix effect		

Egger et al, 1985	28	Overactive/ Atypical Hyperactive	3-12 yrs	yes - in Phase I (FED/Oligoantigenic)	Challenge w one of various food items (milk, oj, wheat, tinned food, 150mg tartrazine, or 150mg benzoic acid)	1 week daily challenge with one incriminated provoking food item	C/Behavior Rating (Parents), Overall Behavior Assessments (Parents, Clinician, Psychologist), and Psychometric Testing (actometer, match familiar figues, Porteus maze); skin prick test for allergy and serum Ig E levels.	*- questionable blinding effectiveness esp colors; subjective mutlitple source findings not confirmed by objective psychological tests*	food challenge effect		food challenge effect	no food effect	no food effect		
David, 1987	24	Hetero-geneous *(6 hyper-active; 18 non-hyperactive w problem behavior)*	5 yrs (2-12)	yes (FED) *(purely behavioral immediate reaction to food additive)*	Drinks w 50 and 250 mg Tartrazine (#5), or 50 and 250 mg Benzoic Acid *(use straw to minimize color detection)*	1 challenge day per treatment (both dosages)	Record of Behavior Changes (Parents, Nurse Staff) - no structured scoring system	*- no randomized treatment order; incomplete blinding; no structured behavioral rating system; no statistical analyses of data*	no tartrazine or benzoic acid challenge effect		no tartrazine or benzoic acid challenge effect				
Rowe, 1988	8	Hetero-geneous *(hype-ractive, non-hyperactive w problem behavior)*	3-15 yrs	yes - in Phase I (F/KP) *(diet responsive and behavior reactive to foods w artificial color)*	Capsules (inner caps) w 50 mg Tartrazine (#5) or 50 mg Carmoisine (tested separately)	Each color daily for 1 week	Behavior Ratings of frequent symptoms (Parents) (attempted Teacher ratings incomplete)	*- unclear blinding effectiveness; single outcome measure (parent rating)*							Parents ratings alone - 2/8 respond to tartrazine & carmoisine; effects not characteristic of hyperactivity (irritab/restless/sleep)
Kaplan et al, 1989	24	Hyperactive w sleep problems	4-6 yrs	not determined prior to study	Elimination Diet (Alberta Children's Hospital Diet)	4-week experimental diet treatment	C w Individualized/ Behavior Rating (Parents, Day Care Worker), Sleep Measures and Physical Symptoms (Parents), Clinical Chem/Hematol and Physical Measures (Clinician) *(Independent observer ratings and Psychometric testing were discontinued)*	*- planned independent observer ratings and objective psychometric testing were not completed; parent findings not confirmed by day-care worker ratings*	FED diet effects		no FED diet effects				Parent ratings and improved sleep records showed 14/24 children with some behavior improvement and sleep improvement on ACS diet but not corroborated by day-care worker behavior ratings

Wilson and Scott, 1989	19	Non-hyperactive	2-13 yrs (11m,8f)	yes (F-color free diet) *(definite history of respiratory, dermatologic, behavioral [4] response to yellow color)*	Drinks w 8.5 mg each of Tartrazine (#5) plus Sunset Yellow (#6), and w 12.5 mg metabisulphite plus 55 mg benzoate	one challenge drinkdaily for 12 days (repeated twice)	Behavioral Assessments: Daily Diary and Overall (Parents) *(unstructured assessment of "relevant" symptoms)*	*- unclear blinding effectiveness; minimal unstructured, unvalidated behavioral rating; single source outcome measure (parent); no statistical analysis of data*							Parent assessment revealed only 1/19 children to show adverse behavioral response to preservative (not color mix) challenge.
Pollock and Warner, 1990	19	Hetero-geneous *(1 hyperactive & 1 retarded)* w problem behaviors	3 - 15 yrs	yes (FCFD) *(problem behaviors of poor concentration and fidgetiness returned after consuming <u>even small amounts</u> of food additives)*	Capsule w 125 mg color mix	two 1-week daily challenges	C/Behavior Rating (Parent daily) and Overall Behavior Assessment (Parent weekly)	*- unclear blinding effectiveness; single source outcome measure (parent rating)*	color mix challenge effect w daily ratings but not weekly overall assessment						
Sarantinos et al, 1990	12	Attention Deficit Disorder (ADD)	4-14 yrs	yes (Color-free Diet) *(9 definite; 4 uncertain)*	Orange juice w 10 mg Tartrazine (#5) or Sunset Yellow (#6)	color challenge on 3 to 6 random days over 28-day period	Conners and Rowe/Behavior Rating (Parents daily)	*- unclear blinding effectiveness; single outcome measure (parent); incomplete information; unknown groups make-up*							Parent ratings show 2/12 ADD children respond to challenge of tartrazine/sunset yellow with irritability, impulsivity, restlessness & sleep problems (questionable reliability)

Egger et al, 1992	40	Hyperactive	9 yrs	yes - in Phase I (FED/oligoantigenic)	Enzyme Potentiated Desensitization (EPD) (intradermal injection w mixed multiple food antigens, additives, color/ preservative mixes) and re-introduction of multiple food items.	3 intradermal injections at 2-monthly intervals and subsequent re-introduction of provoking food items.	Parent daily log of hyperactivity and other symptoms (e.g., headache, recurrent abdominal symptoms). Effectiveness of EPD treatment based on parent stopping reintroduction of provoking food, end-of-trial overall assessment of symptom improvement, and end-of-study Global Assessment.	*- treatment solutions appear not blinded; single outcome measure (parent ratings); limited behavior assessments.*							Based on parental evaluations only, 15/20 food intolerant hyperkinetic subjects receiving EPD desensitization treatment developed tolerance to previously provoking food items (double-blind evidence of food intolerance) , compared with 7/20 placebo treated children.
Carter et al, 1993	19	Hyperactive	3-12 yrs	yes - in Phase I (FED/"few foods" diet)	Challenge with multiple food items and additives (including color capsule w 6.5 mg mixed color; maximum of 26 mg) eliciting adverse rx in open food reintroduction phase (at least 3 food items each child)	food items or placebo given daily for one week (w 2-week washout interval)	C/Behavior Rating (Parent), Global Rating of Severity of Behavior Problems (Parent), Psychologist rating of hyperactive behavior (fidgetiness, restlessness, inattentiveness), Paired associate learning test, and Matching familiar figures (impulsiveness).	*- reasonable verification of blinding but unclear if blinding continued throughout treatment period; probability of false positives with multiple t-testing for only few, if any, effects*	food challenge effect (low magnitude; more irritability than attention deficit)		food challenge effect (pychologist rating w greatest effect on fidgetiness)	food challenge effect on impulsiveness test but not paired associate learning			
Boris and Mandel, 1994	16	Hyperactive	7.5 yrs (11m,5f)	yes - in Phase I (FED/many foods elimination diet)	soup or cranberry sauce w a provoking food item *(5 g powdered food or 100mg color mix: milk, colors, corn, wheat, soy, oranges)*	random 1 to 3 days challenge in one week	C/Behavior Rating (Parents)	*- unclear blinding effectiveness; single outcome measure (parent rating)*	food challenge effect						Graphed data of parental ratings alone show that 11/16 children responded with adverse behaviors to provoking food item challenge.

Rowe and Rowe, 1994	34/20	Problem Behavior w Suspect Hyperactivity / Normal Control	2-6 & 7-14 yrs	yes - in Phase I (color free diet) (23 definite; 11 uncertain)	double-capsule (lactose) or orange juice w 1, 2, 5, 10, 20 and 50 mg Tartrazine (#5)	single challenges w multiple dosages over 21 days	Behavior Rating 30-item Inventory and Conners Behavior 10-item Hyperactivity Rating (Parents), Atopic history, and Skin test for 8 common allergens and for tartrazine.	*- issues of blinding effectiveness and single outcome measure do not negate reliability of findings*							Parent ratings show tartrazine dose related adverse behaviors in 22/34 problem behavior & 2/20 control children (irritab, fidget, sleep probl)
Schmidt et al, 1997	49	Disruptive Behavior Disorders & Hetero-geneous Hyper-activity/ Conduct Disorder	6-12 yrs (47m/2f)	no	Oligoantigenic diet and non-blind methylphenidate (in-patient testing)	oligoantigenic diet treatment for 9 days	3 Behavioral Ratings - teacher in classroom (Conners rating), observers in play and lab testing settings (devised checklists) and 3 Objective Lab Tests (continuous performance test/CPT, paired associate learning/PAT, actometer) (Note: lab test data not used/analyzed)	*- psychometric test data not analyzed or used for assessment*		No oligo diet effect in classroom	Oligo diet effects (in play and lab test settings)	*Test measures taken but not used*	*Actometer measures taken but not used*		12/49 (24%) children responded to oligoantigenic diet on two behavior ratings (in play and lab test settings), but no confirmation by teacher classroom behavior ratings
Uhlig et al, 1997	12	Hyperactive	6-15 yrs	yes - in Phase I (FED-oligoantigenic)	Provoking foods	5 day ingestion of diets with provoking foods	C/Behavior Rating (Investigator during EEG sessions), EEG	*- no blinding used*			provoking foods effects on behavior ratings			EEG-provoking foods effects	
Bateman et al, 2004	277	General Population (non-ADHD) (111 w/166 wo designated hyperactive and 115 w /162 wo atopy)	3 yrs	no	fruit juice w 20 mg color mix and 45 mg sodium benzoate	daily challenge with color/benzoate for 1 week	Aggregated Parental Hyperactivity Ratings/APHR (Parents) and Clinical Aggregated Test Hyperactivity, ATH (Psychologist rated observations and performance tasks)	*- possibly incomplete blinding*	color mix/ benzoate challenge effect (placebo and active rx increased hyper-activity, active signif. more)		no color mix/benzoate effect	no color mix/benzoate effect			

McCann et al, 2007	297	General Population (non-ADHD)	3 yrs (153); 8/9 yrs (144)	no	Fruit juice w 2 color mixes (Mix A and B) @ 20 & 30 mgs (preschool) and 25 & 62 mgs (school age), each w 45 mg sodium benzoate	daily challenge w Mix A and Mix B, each for 1 week	Behavior Ratings (Teacher, Parent, Trained Observer); and Clinical Lab Test(Continuous Performance Task) 8/9 year olds only).Aggregate Global Hyperactivity Score (GHA) used for analyses.	*-inconsistent age/challenge effects; time of challenge uncontrolled*	Component source measure: Mix A effect (not Mix B) in 3 yr olds and Mix B effect (not Mix A) in 8/9 yr olds	Component source measure: no effects of either Mix A or B in either age group	Component source measure: no effects of either Mix A or B in either age group	Component source measure: slight color mix effect in 8/9 yr olds			Composite GHA scores showed 3 yr olds responsive to Mix A but not Mix B and in contrast showed 8/9 yr olds responsive to Mix B but not Mix A.

*** Sex composition of actual diet/challenge study groups was not always presented but did not appear to be a significant factor associated with treatment effects.**
**** Types of diets:(1) *FD = Feingold Diets* - original diet eliminates artificial colors, flavors and natural salicylates; modified diet - eliminates artificial colors, flavors and some preservatives (BHA, BHT, benzoate) but allows salicylates; (2) *FCAFD= Food and Color Additive Free Diet* ; and (3) *FED = Foods Elimination Diets* (also known as oligoantigenic and few foods diet) - eliminates various food items and components, including antigenic elements, artificial colors, flavors and some preservatives**
***** Hyperactive = Hyperkinetic = Attention Deficit Hyperactivity Disorder**

Color Code†:

Grey background	Prior to 1982 NIH Consensus Report. (The trials are listed in chronological order in column 1, from Conners et al (1976) to Spring et al (1981), in grey, then from Salamy et al (1982) to McCann et al (2007).
Pink background	The treatment and treatment duration columns (columns 6 and 7) are color-coded pink for the Feingold/KP Diet (2) or Color Challenge Trials (24).
Blue background	The treatment and treatment duration columns (columns 6 and 7) are color-coded blue for the Elimination Diet (2) and Food Items Challenge Trials (5).
Yellow background	Treatment effects noted, and such treatment effects are described in the table cells.
Green background	No treatment effects found, and such is described in the table cells.
Red text	Notable but non-significant effects or inconsistent findings
Blue text	Cite non-ADHD behavior effects (e.g., sleep abnormality, irritablility, fidgetiness, restlessness) more than attention deficit related behaviors.

†The color-coding is for quick visual reference, but it is not required to access and interpret the data.