0196-206X/00/2506-0423
DEVELOPMENTAL AND BEHAVIORAL PEDIATRICS
Vol. 25, No. 6, December 2004

Printed in U.S.A.

Review Article

Do Artificial Food Colors Promote Hyperactivity in Children with Hyperactive Syndromes? A Meta-Analysis of Double-Blind Placebo-Controlled Trials

DAVID W. SCHAB, M.D., M.P.H.
Columbia University, Department of Psychiatry & The New York State Psychiatric Institute, New York, New York

NHI-HA T. TRINH, M.D., M.P.H.
Harvard University, MGH/McLean Department of Psychiatry, Boston, Massachusetts

ABSTRACT. Burgeoning estimates of the prevalence of childhood attention-deficit/hyperactivity disorder (ADHD) raise the possibility of a widespread risk factor. We seek to assess whether artificial food colorings (AFCs) contribute to the behavioral symptomatology of hyperactive syndromes. We searched ten electronic databases for double-blind placebo-controlled trials evaluating the effects of AFCs. Fifteen trials met the primary inclusion criteria. Meta-analytic modeling determined the overall effect size of AFCs on hyperactivity to be 0.283 (95% CI, 0.079 to 0.488), falling to 0.210 (95% CI, 0.007 to 0.414) when the smallest and lowest quality trials were excluded. Trials screening for responsiveness before enrollment demonstrated the greatest effects. Despite indications of publication bias and other limitations, this study is consistent with accumulating evidence that neurobehavioral toxicity may characterize a variety of widely distributed chemicals. Improvement in the identification of responders is required before strong clinical recommendations can be made. *J Dev Behav Pediatr 25:423–434, 2004.* Index terms: *attention-deficit/hyperactivity disorder, artificial food colorings, complementary and alternative medicine.*

BACKGROUND

Approximately 3–10% of school-age children suffer from attention-deficit/hyperactivity disorder (ADHD).[1–3] Impaired relationships with family and peers, decreased academic achievement, and heightened risk for later drug abuse are some psychosocial costs of this disorder.[1,4] Sufferers have more than double the median health care costs of their peers[5] and may cost school systems an additional 3 billion dollars per year.[1] Common treatments include behavioral therapy and pharmacotherapy with psychostimulants.

Because of concern about the safety of drugs, clinicians and families have for some time been reaching for alternative treatments, such as the Feingold Diet (FD), which eliminates a variety of artificial food colors (AFCs), naturally occurring salicylates, and artificial flavors and particular preservatives. Although surveys have identified widespread use of dietary treatments,[6,7] physicians underestimate the prevalence of such alternative therapies because they overestimate parents' willingness to disclose their use.[8,9]

Controlled trials of the FD and of the effects of substances excluded from the FD first began to be published in 1976.[10,11] Before long, laboratory work began to suggest that azo and fluorescein dyes, which make up the majority of AFCs, may affect motor systems in mammals through dopamine pathways.[12–17]

Many qualitative reviews and one meta-analysis have addressed the effects of either the FD or specific foodstuffs excluded from it, such as AFCs. Qualitative reviews have generally been discordant and contentious,[18–40] though a panel sponsored by the National Institutes of Health in 1982 fashioned a compromise, acknowledging that controlled trials "did indicate a limited positive association between 'the defined diets' and a decrease in hyperactivity."[41]

In their meta-analysis of the effect of the FD on hyperactivity, Kavale and Forness included trials of hyperactive and nonhyperactive children.[42] They folded together trials of the FD, trials of variant diets eliminating a variety of foodstuffs, and trials in which subjects were challenged with individual foodstuffs, including AFCs. Their initial analysis included prospective, retrospective, cross-sectional,

Address for reprints: Dr. David W. Schab, New York State Psychiatric Institute, 1051 Riverside Drive Unit #84, New York, New York, 10032; e-mail: ds2140@columbia.edu.

blind, and nonblind controlled trials that enrolled both hyperactive and nonhyperactive children and employed many categories of outcomes. The authors concluded that placement of children on the FD was associated with a statistically nonsignificant improvement of one-tenth of a standard deviation (SD). In a subanalysis, challenge with substances excluded from the FD (including AFCs) resulted in a statistically nonsignificant worsening of one-twentieth of a SD. The breadth of those authors' inclusion criteria, their oversight of several relevant trials, the subsequent publication of additional relevant trials, and other limitations of their study call for focused consideration of whether AFCs promote symptoms of hyperactivity.

OBJECTIVES

We wish to evaluate whether AFCs contribute to the symptomatology of childhood hyperactivity in children diagnosed with hyperactive syndromes, as measured on behavioral rating scales. Additionally, we propose three subhypotheses whose confirmation would either help explain the heterogeneity of prior trials' results, facilitate identification of responders to AFCs, or aid in design of future trials.

First, we propose that parents and teachers differ in their reports of responsiveness to AFCs, as teachers may be more likely to report hyperactivity than parents in pharmacotherapy trials.[43]

Second, on the grounds that responsive children could ideally be identified without subjecting them to long, complex, blinded trials, we suggest that open trial and parental report can be employed as screening methods to identify potentially responsive children.

Third, we propose that subjects whose diagnosis of hyperactivity was assigned through comprehensive, rigorous evaluation have greater responsiveness to AFCs than do subjects who either were diagnosed more informally or lack a diagnosis of hyperactivity altogether. Our proposition reflects several investigators' suggestion[44–46] that hyperactive and control children fundamentally differ in their physiological responses to AFCs and implies that certainty of diagnosis would be predictive of response to AFCs.

METHODS

Inclusion Criteria

We searched for all double-blind placebo-controlled trials addressing a proposed relationship between consumption of AFCs and behavioral change in children with a diagnosis of hyperactivity. Randomization and employment of any of several reversal designs[47] were acceptable means of allocation compatible with maintenance of double-blind conditions, which reduce the risk of several types of bias.[48]

Trials had to study subjects under 18 years of age and, to be eligible for our primary analysis, enroll children who met diagnostic criteria described at the foot of Table 1. These criteria were chosen as encompassing a spectrum of illness variously labeled over time minimal brain dysfunction, hyperkinesis, hyperkinetic reaction, hyperactivity, attention deficit disorder, and ADHD. Since few investigators actually distinguished between the currently recognized hyperactive, inattentive and combined subtypes of the syndrome, we could not incorporate such a distinction in our study. Subjects may have previously been screened for responsiveness to either AFCs or the FD. A trial would be excluded if its subjects had participated in another included trial. For a secondary analysis aimed at addressing our subhypotheses, we included trials which evaluated whether AFCs provoke hyperactivity symptoms in nonhyperactive children (Table 2).

Trials had to employ an intervention that could isolate the effects of AFCs. Investigators could either 1) challenge hyperactive children with AFCs or placebo or 2) submit subjects to blinded diets with or without AFCs.

We included a wide variety of behavioral outcome measures quantifying hyperactive behavior (Table 1), including all examples of the Conners Parent Teacher Questionnaire (CPTQ),[49] whose 10 items are common to the longer parents' (CPQ) and teachers' (CTQ) scales.[50,51] We preferred Conners' instruments to other behavioral measures because of their widespread employment in research, and we favored the CPTQ in particular because of its known sensitivity to treatment effects in pharmacotherapy trials.[52] A score of 15 on the CPTQ reasonably discriminates hyperactive from control children.[53] Follow-up–the amount of time between ingestion of AFCs or placebo, or between initiation of active or placebo diet and outcome measurement–could be of any length.

Search Strategy & Data Abstraction

We searched through OldMedline, MedLine, PubMed, PsychInfo, Digital Dissertations/UMI ProQuest, ToxLine, Current Contents, Biosis, the Cochrane Controlled Trials Register, and EMBASE with combinations of the following keywords: "hyperactiv#," "hyperkin#," "behav#," "attention deficit," "minimal brain," combined with "color#," "colour#," and "dye." We conducted searches with the names of individual AFCs, examined 5 relevant dissertations with extensive bibliographies,[18,19,54–56] and combed trials and reviews for references to other trials. We contacted investigators whose published work otherwise met our inclusion criteria but lacked sufficient information for inclusion.

We developed a form, available from the authors, for recording trials' features relevant to our hypotheses. Adapting the methods of Alderson et al,[57] we graded trials according to the degree that they adequately addressed types of bias and avoided important threats to validity: differential selection, performance, attrition, detection, and reporting.

Statistical Methods

We employed the DerSimonian and Laird random effects model[58] with the standardized mean difference (SMD) as our 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. Adverse effects on hyperactivity are reflected in positive ES values.

When the SMD could not be calculated directly, we employed conservative statistical techniques to calculate or

Table 1. Characteristics of Trials Included in the Primary Analysis

Author/Year	Diagnostic Criteria[a]	Subjects[b]	Total Days in Active Challenge Arms	Method of Ascertaining Prior Responsiveness	Intervention: AFC, mg/d	Behavioral Outcome (Rater of outcome)
Rose 1977[79,107]	B	2/2	2	Parental report	Tartrazine, 1.2	Author's 3-variable measure[c] (H)
Goyette et al 1978a[78,88]	B, C (15)	16/16	28	Open trial	Mixture, 26	CPTQ (P, T)
Goyette et al 1978b[77,89]	B	13/13	7	Open trial	Mixture, 26	CPTQ (P)
Harley et al 1978[82]	(A: DSM-II, or C), D	9/9	28–42	Earlier trial of the FD	Mixture, 54	CPTQ (P, T)
Levy et al 1978[86]	B	20/20	14	None	Tartrazine, 1	CPTQ (H, P, T)
Levy and Hobbes 1978[94]	B, C (15)	8/8	14	Open trial	Tartrazine, 4	CPTQ (P)
Mattes and Gittelman-Klein 1978[83]	A: DSM-II + CPQ	1/1	10	Parental report & blinded dose-ranging trial	Mixture, 3/5 of "daily average"	CPTQ (P, T)
Rapp 1978[93]	B	24/24	0.5	None	Sublingual mixture	"Global impressions" (P)
Williams et al 1978[84]	B, E	24/24	8	None	Mixture, corresponding to "daily average"	CPTQ (P, T)
Conners 1980[75]	C (15)	12/30	14	None	Mixture, 26	CPTQ (P)
Swanson and Kinsbourne 1980[80]	B, E	20/40	1	None	Mixture, 100–150	CTQ (H)
Adams 1981[95]	C (15), D	18/18	1	Parental report	Mixture, 26.3	"Parental report" (P) & author's 2-variable measure[d] (H)
Spring et al 1981[81]	C (15), D	6/6	6	Parental report	Mixture, 26	CPTQ-modified (P, T)
Sarantinos et al 1990[85]	A: DSMIII-R	12/12	6	Parental report for 9/13 of original subjects	3 Days tartrazine, 10 & 3 days sunset yellow, 10; or 6 days tartrazine, 10	CPTQ (P), RBRI (P)
Rowe and Rowe 1994[76]	D	34/54	6	Open trial	Tartrazine, 1, 2, 5, 10	RBRI (P)

P, parents; T, teachers; H, health care providers; FD, Feingold Diet; AFC mg/d, total milligrams of artificial food coloring per day; CPTQ, Conners Parent Teacher Questionnaire;[49] RBRI, Rowe Behavior Research Inventory;[111] CTQ, Conners Teacher Questionnaire.[92]
All trials employed double-blind, crossover design; all employed randomization except Rose 1977.
[a]Diagnostic criteria key: **A.** Diagnosis with reference to DSM-III[108] criteria for ADD with or without mention of hyperactivity; or to DSM-IIIR[109] criteria for ADHD or undifferentiated ADD; or to DSM-II[110] criteria for Hyperkinetic Reaction of Childhood if one of the Conners scales was also used; **B.** Diagnosis of "hyperactivity" clearly based on clinician's evaluation but without reference to aforementioned criteria; **C.** CPTQ cutoff (score), which, if without mention of criteria A, B, D, or E, is the sole criterion for inclusion here; **D.** "Hyperactive" or "referral" for hyperactivity without further details; **E.** Responsiveness to medication; **F.** Heterogeneous; **G.** Nonhyperactive.
[b]Subjects who contributed data to the primary analysis/total subjects in the trial.
[c]Author operationalized 3 variables: duration the subject spent attending to an assigned task; frequency/duration spent out-of-seat in classroom; and frequency of acts of physical aggression.

estimate it.[59–61] To calculate the variance around the SMD for crossover trials, we employed a method, described in a recent meta-analysis,[62] involving the Pearson product-moment correlation (r) to reflect correlation between individuals' scores in crossover periods. When trials provided insufficient information by which to calculate r, we imputed a value from other trials through standard techniques of combining r values.[63] For trials reporting noncontinuous (e.g., dichotomous) outcomes, we converted these data to the most conservative odds ratio that would fit the data. We then calculated the SMD through methods described by Shadish et al[61] and the variance through a method described by Rosenthal.[64] Our subhypotheses, which propose differences between groups of studies, were tested with weighted analysis of variance, which weights trials' ESs by the inverse of their variance.[65]

When the Mantel-Haenszel test of homogeneity demonstrated that within-group statistical homogeneity was unlikely ($p < 0.20$), we defined ESs as statistically heterogeneous and therefore not combinable.[66,67] We conducted additional analyses to test the sensitivity of our methods to several assumptions. A funnel plot was created to evaluate the potential for publication bias. The robustness of the results against publication bias was estimated with the fail-safe n. This statistic represents the number of unpublished studies with an ES of zero necessary to reduce a summary ES to a particular value;[68] it assumes the unpublished studies are of similar size to those already published. To calculate our fail-safe n, we chose an ES of 0.15, which, even if statistically significant, would represent a limited threat to an individual's health.

RESULTS

Search Results

Of 2156 references found by electronic search through September, 2002, 427 nonduplicate items were identified as

Table 2. Characteristics of Additional Trials and Subgroups Considered in the Secondary Analysis

Author/Year	Diagnostic Criteria[a]	Subjects[b]	Total Days in Active Challenge Arms	Method of Ascertaining Prior Responsiveness	Intervention: AFC, mg/d	Behavioral Outcome (Rater of Outcome)
Conners 1980-s[75]	G (CPTQ <15)	18/30	14	None	Mixture, 26	CPTQ (P)
Weiss et al 1980[69]	G	22/22	1	Parental report	Mixture, 35.26	CPTQ (P) & measures chosen by parents (P)
Mattes and Gittelman 1981[70]	A (DSM-III): 5 of 11, otherwise F	11/11	14	Parental report & blinded trial	Mixture, 78	CPTQ (P, T); HA items from CDS (H)
Thorley 1984[71]	F, all mentally retarded	10/10	2	None	Mixture, 91.8	CTQ (T); HA factor of the CPQ (P)
David 1987[72]	A (DSM-III): 6/24, otherwise D	24/24	1	Parental report	Tartrazine, 300	"Global judgment" (P, H)
Rowe 1988[73]	F	8/8	28	Open trial	Carmoisine, 50 & tartrazine, 50 separately	Author's 8-item behavioral scale (P)
Pollock and Warner 1990[74]	F	19/19	2	Parental report	Mixture, 125	CPTQ (P)
Rowe and Rowe 1994-s[76]	G	20/54	6	None	Tartrazine, 1, 2, 5, 10, 20, & 50 separately	30-Item RBRI (P)

P, parents; T, teachers; H, health care providers; CDS, Children's Diagnostic Scale[49]; CPTQ, Conners Parent-Teacher Questionnaire[49]; CTQ, Conners Teacher Questionnaire[92]; CPQ, Conners Parent Questionnaire[49]; HA, hyperactivity; AFC mg/d, total milligrams of artificial food coloring per day; the *s* suffix after several trials indicates that the table reflects data excluding a subset of patients whose data is reflected in Table 1. All trials employed randomized, double-blind crossover design.

[a]Diagnostic criteria key: **A.** Clinician's diagnosis with reference to DSM-III[108] criteria for ADD with or without hyperactivity; or to DSM-IIIR[109] criteria for ADHD or undifferentiated ADD; **B.** Diagnosis of "hyperactivity" clearly based on clinician's evaluation but without reference to aforementioned criteria; **C.** CPTQ cutoff (score), which, if without mention of criteria A, B, D, or E, is the sole criterion; **D.** "Hyperactive" or "referral" for hyperactivity without further details; **E.** Responsiveness to medication; **F.** Heterogeneous; **G.** Nonhyperactive.

[b]Subjects who contributed data to the secondary analysis/total subjects in the trial.

pertinent trials or as potentially containing references to such trials. For our primary analysis, we identified 15 unique double-blind placebo-controlled trials evaluating the behavioral effects of AFCs among subjects whose baseline diagnosis of hyperactivity has been graded (Table 1). The grading roughly reflects the comprehensiveness and rigor of the diagnostic process employed by the trials' investigators. For our secondary analysis, we identified eight additional crossover trials or subsets of trials which otherwise met our inclusion criteria but studied either exclusively nonhyperactive children[69] or heterogeneous groups of children (Table 2).[70–76] Two of the latter,[75,76] enrolling both hyperactive and nonhyperactive children, contribute independent data to both the primary and secondary analyses. A table of excluded trials is available from the authors.

Primary Analysis

Fifteen trials included a total of 219 subjects (Table 1). The average age of the subjects was estimated at 7.9 years and the male-to-female ratio at 5.5 to 1. Seven trials explicitly excluded children on medications or discontinued the medications before the beginning of the trial.[75,77–82] Of the 219 participants, 136 (62%) entered the double-blind phase of the trials after successful screening for responsiveness to either AFCs or the FD.

All trials were double-blind crossover trials. All trials employed randomization except one,[79] which fixed order of treatment beforehand. Only 5 trials had washout periods, all between 2 and 5 days.[76,81,83–85] Five trials[78,82–84,86] employing the CPTQ provided insufficient information whereby to calculate r for at least 1 set of raters; we imputed a value (0.822) through combination of data from similar relevant trials.[74,75,77,78,87–89] This value is consistent with published values as employed in pharmacotherapy trials.[51,90–92] Three trials reported outcomes that were dichotomous[76,85] or trichotomous.[93]

Tartrazine and a variety of mixtures of AFCs were the primary interventions in the challenge arms of the trials and were administered through a variety of vectors, such as pills and cookies. Dosages varied widely and, in 8 trials,[75,77,78,81–84,94] were split into multiple daily doses. Three trials employed outcome measures that have not been validated.[79,81,95]

Only 2 trials[82,84] received our highest validity score, "A." Rose's trial[79] received a "C" for failure to employ randomization, as described; Rapp's trial[93] received a "C" for employing what appeared to us a poorly disguised placebo. Failure to discuss procedures of allocation concealment was the most common mark against the trials.

Secondary Analysis

Eight crossover trials included 132 participants (Table 2). The male-to-female ratio was approximately 3.5 and the average age of participants was 7.3. Two of these trials contributed hyperactive patients to our primary analysis.[75,76] Eighty-four patients (64%) in these trials had been deemed responsive by screening before entry into the blinded trial. Two trials had no wash-out period.[69,75] Three

trials employed unorthodox outcome measures.[69,72,76] One[72] of these received the only rating of "C," assigned for imperfectly satisfactory blinding. No trial received a score of "A."

Results

Primary Analysis. The trials' summary ES is 0.283 (95% CI, 0.079 to 0.488), reflecting a change of slightly more than one quarter of a SD (Figure 1). If we assume a normal distribution of response to AFCs, this change represents a shift from the 50th to the 61st percentile of hyperactivity for the average hyperactive child in the population of trials.

To address our first subhypothesis, we have separated parents', teachers', and clinicians' ratings into 23 ESs (Figure 2). While health professionals' ratings (ES = 0.107, 95% CI, −0.128 to 0.343) and teachers' ratings (ES = 0.0810, 95% CI, −0.073 to 0.235) are not statistically significant, parents' ratings are (ES = 0.441, 95% CI, 0.161 to 0.721), corresponding to a 17 percentile shift for the average hyperactive child. Despite an apparent reduction in the likelihood of statistical heterogeneity within the groups (p = 0.506 vs. 0.232 in the unstratified analysis), no statistical difference was found between these three groups' scores (p = 0.674).

To address our second subhypothesis, we compared the group of trials whose subjects had not been screened for responsiveness to AFCs or the FD to the group of trials whose subjects had been screened by either an open trial or parental report. The summary ES of the trials without a criterion of responsiveness demonstrated a statistically nonsignificant ES of 0.090 (95% CI, −0.108 to 0.288). The summary ES of the trials whose subjects had previously demonstrated responsiveness through open trial or parental report had a statistically significant ES of 0.535 (95% CI, 0.149 to 0.920), corresponding to a 20 percentile shift for the average nonhyperactive child screened for responsiveness to AFCs. An additional trial, whose subjects had been screened with a blinded trial of the FD, had a statistically nonsignificant ES of 0.182 (95% CI −.208 to 0.571). Segregation of the trials into three strata, according to their inclusion criteria, appears to reduce within-group statistical heterogeneity (p = 0.448 versus prior p = 0.232), suggesting a more natural grouping of studies. However comparison of the group without a screening criterion to the group that screened through open trial or parental report falls short (p = 0.185) of substantiating a statistically significant difference between these groups.

To begin addressing our third subhypothesis, we ordered trials by the diagnostic grades found in Table 1. We found no visual trend among their ESs (Figure 1).

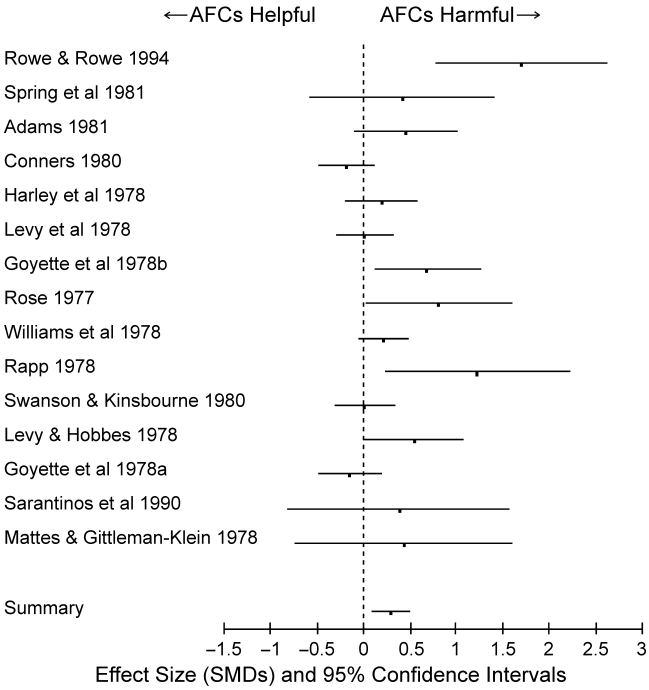

FIGURE 1. The effect sizes from the 15 trials in the primary analysis and their resultant summary effect size. Trials are listed in an order broadly reflecting the comprehensiveness and rigor employed in the assignment of subjects' baseline diagnoses: the diagnostic processes of trials listed further down the page received higher marks, according to the criteria described at the foot of Table 1. AFC, artificial food color; SMD, standardized mean difference.

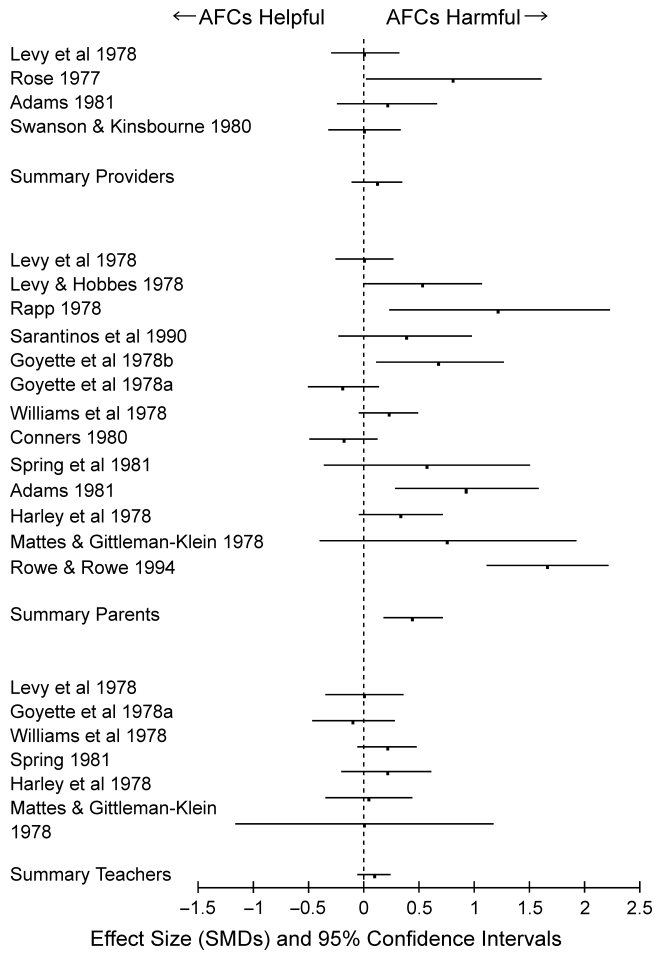

FIGURE 2. The effect sizes from the 15 primary trials (or subgroups) are stratified by clinicians', parents', and teachers' ratings. AFC, artificial food color; SMD, standardized mean difference.

Secondary Analysis. For the secondary trials, our summary ES of 0.117 (95% CI, −0.113 to 0.347) was not statistically significant. Trials that did not screen for responsiveness demonstrated a statistically nonsignificant response to AFCs (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 screen through open trial or parental report (Figure 3). This ES corresponds to a 12 percentile rank shift in hyperactivity for the average nonhyperactive average child among trials that screened for responsiveness prior to double-blind allocation. The difference in the ESs of screened and unscreened children in the secondary studies is statistically significant (p = 0.022), supporting our second subhypothesis. A post-hoc analysis combining trials from both the primary and secondary analyses suggested that, among all trials, those that screened with open trial or parental report differed significantly from those that did not screen (p = 0.022).

The summary ES of the trials in the secondary analysis, whose participants were predominantly nonhyperactive, did not differ significantly from the summary ES of the trials in the primary analysis, whose participants all had a putative diagnosis of hyperactivity (p = 0.660). Both this fact and the lack of any visual trend in Figure 1, described above, defy our third subhypothesis that presence of a baseline diagnosis of hyperactivity and rigor of that diagnosis correspond to reactivity to AFCs.

Sensitivity Analysis of Primary Trials. Excluding from our 15 primary trials the 2 trials that received our lowest validity score "C" resulted in a smaller but still statistically significant result (ES 0.216, 95% CI, 0.015 to 0.410). Because the random effects model disproportionately gives small trials more weight, we removed the 2 smallest trials, which enrolled only one[83] or two[79] subjects, on the grounds that their positive ESs may have skewed our overall results. The summary ES was reduced to 0.255 (95% CI, 0.043 to 0.467), but the hypothesis of statistical homogeneity was rejected (p = 0.179) in this one analysis. However, when we implemented both prior conservative manipulations, the assumption of homogeneity was restored, achieving a result that is smaller but of similar magnitude and statistically significant (cf. Table 3).

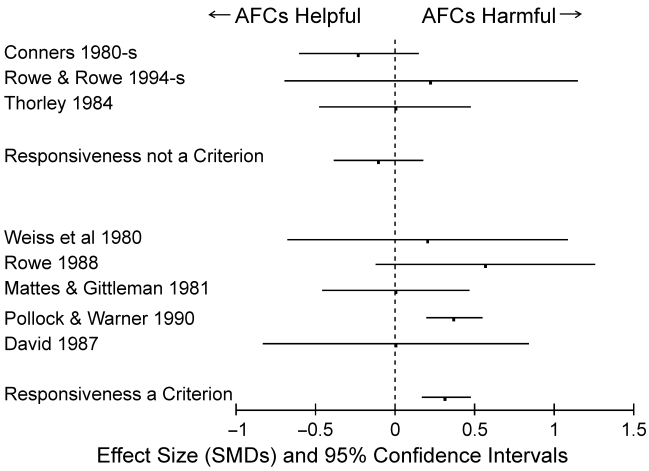

FIGURE 3. The effect sizes from the eight trials (or subgroups of trials) enrolling exclusively nonhyperactive children or children with heterogeneous conditions in the secondary analysis. The suffix *–s* after several trials indicates that the data of nonhyperactive children come from a trial enrolling both nonhyperactive and hyperactive children. AFC, artificial food color; SMD, standardized mean difference.

Because of concern that imputation of too large a value of the product-moment correlation coefficient (r) would reduce trials' variance and therefore artificially narrow their confidence intervals, we retested our primary model by reducing the coefficient to 0.500 in the five trials that required imputation of the coefficient. This maneuver raised the ES to 0.371 with only a slight widening of the confidence interval (95% CI, 0.0953 to 0.647).

The absence of trials from the lower left-hand corner of our funnel plot (Figure 4) corresponds to possible publication bias against small trials with negative ESs. Given a calculated summary ES of 0.283, the fail-safe n for an ES of 0.15 is 13.30, which is almost as many studies as were actually published.

DISCUSSION

Our meta-analysis supports the hypothesis that AFCs promote hyperactivity in hyperactive children, as measured on behavioral rating scales. This, our central result, would appear to conflict with a relevant subanalysis in a prior meta-analytic evaluation of elimination diets. In that subanalysis, Kavale and Forness[42] examined trials challenging subjects with foodstuffs excluded from the FD. Their calculated ES was only 0.045 (95% CI, −0.046 to 0.136), in contrast to our larger ES of 0.283 (95% CI, −0.079 to 0.488). However, in their analysis, the evaluated interventions were not limited to AFCs but also included other substances which may have different, if any, effects on hyperactivity. Additionally, our primary analysis included four trials that theirs did not: two trials that they overlooked[79,95] and two trials published subsequent to their analysis.[76,85] Three of these trials were large, increasing the power of our meta-analysis. Additionally, we enhanced our power by employing the statistical technique of Jennings et al[62] that acknowledges the reduction of statistical variance in crossover trials.

In our sensitivity analysis, removing trials of questionable validity and challenge of our statistical assumptions did not appreciably change the magnitude of the summary ES or its statistical significance. The magnitude of this ES is approximately a third to a half the magnitude of the ESs calculated in a meta-analysis of trials evaluating methylphenidate as a treatment of ADHD.[96] Therefore, the extent of behavioral deterioration posed by AFCs may be understood as a third to a half of what might be observed when hyperactive children are taken off their psychostimulants.

Our first subhypothesis, that parents, teachers and clinicians differ in their reports of hyperactivity, could not be confirmed: the three groups' individual ESs did not statistically differ. Although the summary ES of each of the three groups was positive, the finding that only the parents'

Table 3. Summary of Statistical Analysis

Analysis	Number of Outcomes	Summary SMD	95% CI (−)	95% CI (+)	Test of Heterogeneity Q Statistic	DF	*p*
Primary trials: summary	15	0.283	0.079	0.488	17.47	14	.232
Stratified by rater							
Care staff	4	0.107	−0.128	0.343	3.41	3	.332
Parents	13	0.441	0.161	0.721	13.51	12	.333
Teachers	6	0.081	−0.073	0.235	2.32	5	.803
					Q_w: 19.25	20	.506
Stratified by inclusion criteria							
Prior responsiveness not a criterion	6	0.090	−0.108	0.288	6.14	5	.293
Prior responsiveness by open trial or parental report	8	0.535	0.149	0.920	5.83	7	.560
Prior responsiveness by blind FD trial	1	0.182	−.208	0.571	0.00	0	1.00[a]
					Q_w: 11.97	12	.448
Sensitivity analysis of primary trials							
Excluding validity rating "C"	13	0.216	0.015	0.410	14.99	12	.242
Excluding 2 smallest trials	13	0.255	0.043	0.467	16.28	12	.179
Prior 2 criteria applied together	12	0.210	0.007	0.414	14.52	11	.205
Imputed *r* replaced by 0.500	15	0.371	0.095	0.647	13.17	14	.513
Secondary trials: summary	8	0.117	−0.113	0.347	4.90	7	.672
Stratified by inclusion criteria							
Prior responsiveness not a criterion	3	−0.112	−0.393	0.169	1.12	2	.571
Prior responsiveness by open trial or parental report	5	0.316	0.157	0.475	3.20	4	.525
					Q_w: 4.32	6	.633

SMD, standardized mean difference; CI, confidence interval; DF, degrees of freedom for *within*-group Q; Q_w, Q total for preceding strata; FD, Feingold Diet.
[a]Chi-distribution is not defined for zero DF, but probability of statistical heterogeneity in a stratum of one trial is 0.00 ($p = 1.00$).

summary ES was statistically significant deserves explanation. Because crossover trials enable subjects to compare the two phases, unblinding on the part of the parents, who often administer the intervention, could explain these results. However, all investigators who rigorously tested the blinding[77–79,82,83] confirmed that the subjects, their parents or the investigators were unable to guess beyond chance the identity of the interventions.

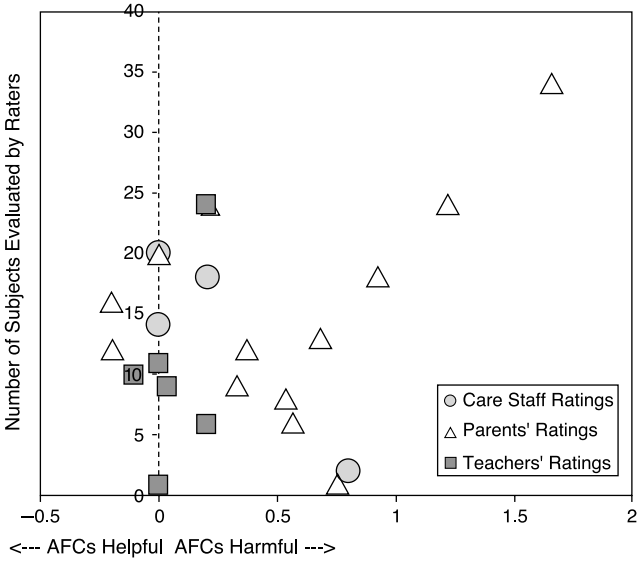

FIGURE 4. Funnel plot of effect sizes recorded by clinicians, parents, and teachers. AFC, artificial food color.

The characteristics of particular rating scales may be the prisms through which to understand parents' penchant for higher scoring in our trials. First, one trial,[76] whose ES was the largest of those we calculated, measured outcome with a scale reflecting parental reports of symptoms suspected to result from AFC ingestion. This scale, the Rowe Behavioral Research Inventory, emphasizes irritability and sleeplessness, both of which may be particularly disruptive to the parent-child dyad, and de-emphasizes restlessness and attentional difficulties, which are of particular concern to teachers. Second, Weiss et al[69] reported a very small probability ($p = 1.6 \times 10^{-5}$) that chance accounted for one particular child's pattern of responsiveness to AFCs. Their trial employed a different outcome scale for each subject tailored to symptoms preselected by the child's parents from standardized rating scales. Third, Mattes and Gittelman-Klein,[83] who found no change on the CPTQ in their single-subject multiple-crossover design, conceded that this scale did not capture the irritability consistently noted by the subject's mother while the subject was receiving the active intervention. These examples suggest several reasons why parents may have detected behavioral change unnoticed by teachers and clinicians: parents' concerns may differ from those of teachers and clinicians; parents may be particularly attuned to the idiosyncrasies of their own children; and AFCs may promote a pattern of symptoms that is incongruent with modern criteria for ADHD but that is nevertheless bothersome to parents.

Our second subhypothesis, that children potentially responsive to AFCs can be identified through screening methods, was partially substantiated by our analysis. As a

group, trials that screened for responsiveness demonstrated a statistically significant effect whether enrolling hyperactive or heterogeneously diagnosed, predominantly nonhyperactive subjects. However, we found a statistically significant difference between trials that screened and those that did not only among the group of trials enrolling heterogeneous subjects predominantly without hyperactivity. 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.

Our third subhypothesis could not be confirmed. There was no statistically significant difference between the ESs of trials enrolling hyperactive children and the ESs of trials enrolling predominantly or exclusively nonhyperactive children. Even among the group of trials enrolling hyperactive children, rigor of diagnosis did not correspond to reactivity to AFCs, suggesting that AFC-sensitivity among patients may not be limited to those with clear-cut, criteria-specific hyperactive syndromes. That we were able to find an overall statistically significant ES for our primary analysis but not for our secondary analysis generally corresponds to evidence from prior studies demonstrating that hyperactive children have physiologic responses to AFCs not found in controls.[44–46]

Most trials evaluated in this meta-analysis employed methods and designs that may potentially underestimate the effect of AFCs. First, the crossover design is appropriate only when effects do not carry over between periods. Rowe found that the effect of AFC challenges carried over in rating scales for 3–4 days in one child, while in another child 3.5 weeks,[73] suggesting that the washout periods used in most trials may have been inadequate. Second, several investigators have demonstrated that objective measures of learning and attention are adversely affected within two hours of AFC administration, suggesting that too much time may have elapsed between administration of AFCs and measurement of outcomes.[23,80,97] Third, as several investigators evaluated dosages well below children's true likely daily exposure,[98] one must consider whether the real-world effects of AFCs are greater than the effects captured in our trials.

Several limitations, especially of the included trials, but also of our meta-analytic methods, restrict the interpretation of our results. First, the use of unorthodox scales to measure behavioral responses may limit the validity of conclusions drawn about the effects of AFCs on hyperactivity. However, while the standard CPTQ was developed as a simplified measure of childhood behavior suggesting a diagnosis of hyperactivity, its specificity in distinguishing hyperactivity from global psychopathology has also been questioned.[52] The likelihood that the multiple versions of Conners' instruments measure slightly different underlying behavioral constructs[52] further challenges the specificity of our findings. Consequently, the trials included in the present study may effectively implicate AFCs more in provoking general behavioral disturbance than in exacerbating criteria-specific hyperactive symptomatology.

Second, many of the trials employ statistical methods that do not distinguish between normally distributed reactivity and large idiosyncratic reactions confined to a few individuals. Several investigators noted that positive group effects are often attributable to the profound response of a few individuals, while conclusions about a lack of group effects may overlook considerable individual response.[34,69]

Third, our funnel plot raises the possibility of publication bias against small trials with negative results. However, the fail-safe n demonstrates that discovery of a considerable number of unpublished studies would be required to reduce our summary ES to a trivial value. Confidence in our overall result is further suggested by our sensitivity analysis, which demonstrates that our results are internally robust.

Finally, clinical heterogeneity may limit the interpretability of our results. We attempted to address clinical heterogeneity by segregating trials which screened for responsiveness to AFCs from those that did not by differentiating between trials which enrolled hyperactive and heterogeneous populations, and by considering whether formal assignment of diagnoses predicted responsiveness. Although tests of statistical heterogeneity suggest that results were statistically combinable, a larger number of trials, especially with more detailed data, would have given us greater power to aggressively address many clinical differences in trials, such as types and dosages of AFCs, timing and choice of outcome measurements, and enrollment of the various subtypes within the hyperactivity spectrum. Trials' inconsistent reporting of subtypes prevents us from retiring an important question: whether the trials' and our study's focus on behavioral outcomes makes our study irrelevant to ADHD's inattentive subtype, whose symptoms may not be predominantly behavioral phenomena.

Despite these limitations, our results strongly suggest an association between ingestion of AFCs and hyperactivity. We believe this is the first comprehensive quantitative analysis of trials addressing the effects of AFCs on hyperactivity. Prior reviews and the only meta-analysis in field of elimination diets have attempted to address this issue but our critical review differs in several ways. We have focused on AFCs in particular rather than the FD as a whole. Compared to the prior meta-analysis by Kavale and Forness,[42] ours employs hypotheses that are more explicit, inclusion criteria that are more rigorous, and a search that is now more current and also more comprehensive. Consequently, we have included only double-blind placebo-controlled trials. Our primary and secondary analyses include 2 trials[79,95] that Kavale and Forness overlooked and 6 trials[71–74,76,85] subsequent to their analysis. In addition, as noted, our statistical techniques more richly exploit the advantages of crossover trials. Furthermore, our more explicit sensitivity analysis and our evaluation of publication bias afford a previously unwarrantable confidence: that any conclusions drawn about the relevant body of literature do not unduly depend on the influence of a few unrepresentative trials. Finally, our evaluation of our subhypotheses has identified areas to be examined and methods to be employed in future research.

Closer characterization of the response to AFCs than we currently have requires studying responders. We have demonstrated that screening through parental report and open trial may aid in identifying such responders. Other

research has hinted that electroencephalography[99] and measurement of simple physiologic parameters[44] could aid in the identification of these responders. Combining these tools could focus investigators' efforts on a more targeted population and could reduce the number of non-responders subjected to long, complex blinded trials. However, we have also shown that researchers should not limit their search for responders to those previously diagnosed with hyperactivity.

In our results, parents tended to detect greater change from the intervention than did teachers. This tendency contrasts that of pharmacotherapy trials, in which teachers' ratings have generally been higher, probably because common dosing schedules optimize children's behavior during the school day.[96] Future research on AFC-induced behavioral disturbance should acknowledge how timing of interventions may affect the ratings of parents, teachers and clinicians. Research should maximize these groups' attunement to particular symptoms, as was done by Weiss et al, whose scales reflect parental concerns.[69] Finally, we advocate for the continuing development of scales which address the particular symptoms that AFCs may cause, such as the insomnia and irritability noted by several investigators.[73,76] Such research may lead investigators out of the realm of hyperactivity and into a wider province of behavioral pathology.

We also recommend that crossover designs be avoided until the pharmacokinetics of AFCs are better understood. The effect of interactions of AFCs with foods and medications should also to be explored. Finally, we recommend that future research avoid the pitfalls of many prior trials by explicitly identifying subjects' demographic characteristics; by employing specific diagnostic criteria and by identifying diagnostic subtypes and comorbidities; by identifying the concurrent use of medications; by specifying the interval between administration of AFCs and measurement of effect; by administering specific AFCs rather than mixtures; by explicitly testing the blinding of subjects; and by reporting all such information in published reports.

Given the need for additional research, we are cautious about making clinical recommendations at this time. On the one hand, parental preference for nonpharmacologic treatment[100] must be acknowledged and, in many cases, accommodated, if the clinician's relationship with a patient's parents is to be preserved.[9] On the other hand, the restrictiveness of an AFC-free diet may burden hyperactive children, who are already at risk for poor psychosocial outcomes.[101] Therefore, imposition of the diet should be done reluctantly until more certain methods have been developed to identify who is AFC-responsive.

Basic questions about the biology of AFCs remain unanswered, including whether children's responses to AFCs depend more on an allergic or a pharmacological mechanism.[99,102,103] Tartrazine and its metabolites, for example, may act through both.[102] Clarification of the mechanism of response will deepen understanding of the possibility that responses are not normally distributed.

Neurochemical research into the basis of ADHD has strongly implicated defects in dopamine transmission. Both dopamine depletion and administration of AFCs create hyperactivity in developing rats.[13] However, a theory on the contribution of AFCs to ADHD symptomatology must contend with two incongruities. First, the symptomatology of ADHD may differ from the pattern of symptoms induced by AFCs: as suggested by Rowe and Rowe,[76] AFCs are associated more with irritability and insomnia than restlessness and inattention. Second, the patterns of behavior elicited by dopamine depletion in developing rats differs from the pattern elicited by administration of AFCs (or their metabolites). The sensitivity of developing rats also begs consideration of whether exposure to AFCs differentially affects the developing and developed organism.

Given how little we know about the neurochemical effects of AFCs, we endorse prior recommendations that assessment of behavioral toxicity should be a part of food additive evaluation.[33] The increasing recognition of subclinical neurobehavioral toxicity of low doses of several environmental toxins, especially among developing children,[104–106] demands ambitious vigil against avoidable harmful exposures. At the very least, regulators should track consumption of AFCs; we know only that domestic production of food dyes quadrupled between 1955 and 1998.[25] Finally, as long as we remain uncertain about the early and long-term effects of these exposures, society should engage in a broader discussion about whether the aesthetic and commercial rationale for the use of AFCs is justified.[107]

REFERENCES

1. National Institutes of Health. Diagnosis and Treatment of Attention Deficit Hyperactivity Disorder: NIH Consensus Statement; Nov 16–18, 1998;16:1–37.
2. Rowland AS, Umbach DM, Stallone L, Naftel AJ, Bohlig EM, Sandler DP. Prevalence of medication treatment for attention deficit-hyperactivity disorder among elementary school children in Johnston County, North Carolina. *Am J Public Health.* 2002;92:231–234.
3. Green M, Wong M, Taylor J, Feinleib M. *Diagnosis of Attention-Deficit/Hyperactivity Disorder.* Rockville, MD: Agency for Health Care Policy and Research, US Department of Health and Human Services; 1999. AHCPR Publication 99-0050.
4. Richters JE, Arnold LE, Jenson PS, et al. NIMH Collaborative Multisite Multimodal Treatment of Children with ADHD. I. Background and Rationale. *J Am Acad Child Adolesc Psychiatry.* 1995;34:987–1000.
5. Leibson CL, Katusic SK, Barbaresi WJ, Ransom J, O'Brien PC. Use and costs of medical care for children and adolescents with and without attenttion-deficit/hyperactivity disorder. *JAMA.* 2001; 285:60–66.
6. Bennett FC, Sherman R. Management of childood "hyperactivity" by primary care physicians. *J Dev Behav Pediatr.* 1983;4:88–93.
7. Stubberfield T, Parry T. Utilization of alternative therapies in attention-deficit hyperactivity disorder. *J Paediatr Child Health.* 1999;35: 450–453.

8. Sawni-Sikand A, Schubiner H, Thomas RL. Use of complementary/alternative therapies among children in primary care pediatrics. *Ambul Pediatr.* 2002;2:99–103.
9. Chan E. The role of complimentary and alternative medicine in attention-deficit hyperactivity disorder. *J Dev Behav Pediatr.* 2002; 23(suppl):S37–S45.
10. Conners CK, Goyette CH, Southwick DA. Food additives and hyperkinesis: preliminary report of a double-blind crossover experiment. *Psychopharmacol Bull*. 1976;12:10–11.
11. Conners CK, Goyette CH, Southwick DA, Lees JM, Andrulonis PA. Food additives and hyperkinesis: a controlled double-blind experiment. *Pediatrics.* 1976;58:154–166.
12. Silbergeld EK, Anderson SM. Artificial food colors and childhood behavior disorders. *Bulletin of the New York Academy of Medicine: Journal of Urban Health.* 1982;58:275–295.
13. Goldenring JR, Batter DK, Shaywitz BA. Sulfanilic acid: behavioral change related to azo food dyes in developing rats. *Neurobehav Toxicol Teratol.* 1982;4:43–49.
14. Logan W, Swanson JM. Erythrosine B inhibition of neurotransmitter accumulation by rat brain homogenate. *Science.* 1979;206: 363–364.
15. Mailman RB, Ferris RM, Tang FL, et al. Erythrosine (Red No. 3) and its nonspecific biochemical actions: what relation to behavioral changes? *Science*. 1980;207:535–537.
16. Mailman RB, Lewis MH. Food additives and childhood hyperactivity. *Bol Asoc Med P R.* 1983;75:551–553.
17. Shaywitz BA, Goldenring JR, Wool RS. Effects of chronic administration of food colorings on activity levels and cognitive performance in developing rat pups treated with 6-hydroxydopamine. *Neurobehav Toxicol Teratol.* 1979;1:41–47.
18. Turman MW. *A Study of Research on the Effects of Food Additives on Children's Behavior as Evidenced by the Hyperactivity/Learning Disabilities Syndrome* [dissertation]. San Diego, CA, United States International University; 1985.
19. Dortch AL. *A Compilation of the Literature Relating to Hyperactivity - Hyperkinesis, Learning Disabilities and Food Additives* [dissertation]. Saint Louis, MO, Saint Louis University; 1980.
20. Breakey J. The role of diet and behaviour in childhood. *J Paediatr Child Health.* 1997;33:190–194.
21. Brenner A. Food additives and behavior. *Md Med J.* 1986;35: 344–345.
22. Brunner RL, Vorhees CV, Butcher RE. Food colors and behavior. *Science.* 1981;212:578–579.
23. Conners CK. *Food Additives and Hyperactive Children*. New York: Plenum Press; 1980.
24. Conners CK. Aritificial colors in the diet and disruptive behavior: Current status of research. In: Knights R, Bakker D, eds. *Treatment of Hyperactive and Learning Disordered Children: Current Research*. Baltimore, MD: University Press; 1980:113–120.
25. Jacobson MF, Schardt MS. Diet, ADHD & Behavior: A Quarter-Century Review. *Center for Science in the Public Interest*. Available at: www.cspinet.org. Accessed December 4, 2000.
26. Rimland B. The Feingold diet: An assessment of the reviews by Mattes, by Kavale and Forness and others. *J Orthomolecular Psychiatry.* 1984;13:45–49.
27. Rippere V. Food additives and hyperactive children: a critique of Conners. *Br J Clin Psychol*. 1983;22(pt 1):19–32.
28. Silver LB. The "magic cure": A review of the current controversial approaches for treating learning disabilities. *J Learn Disabil.* 1987;20: 498–504, 512.
29. Stare FJ, Whelan EM, Sheridan M. Diet and hyperactivity: is there a relationship. *Pediatrics.* 1980;66:521–525.
30. The National Advisory Committee on Hyperkinesis and Food Additives. *Final Report to the Nutrition Foundation.* New York: Nutrition Foundation; 1980.
31. Weiss B. In rebuttal. Food additives and hyperkinesis. *Am J Dis Child.* 1980;134:1126–1128.
32. Weiss B. Food colors and behavior: A response. *Science.* 1981; 212:579.
33. Weiss B. Food additives and environmental chemicals as sources of childhood behavior disorders. *J Am Acad Child Psychiatry.* 1982;21: 144–152.
34. Weiss B. Food additives as a source of behavioral disturbances in children. *Neurotoxicology.* 1986;7:197–208.
35. Wender EH. The food additive-free diet in the treatment of behavior disorders: a review. *J Dev Behav Pediatr.* 1986;7:35–42.
36. Harley JP, Matthews CG. Food Additives and Hyperactivity in Children. In: Knights R, Bakker D, eds. *Treatment of Hyperactive and Learning Disordered Children: Current Research*. Baltimore, MD: University Press; 1980:121–130.
37. Williams JI, Cram DM. Diet in the management of hyperkinesis: a review of the tests of Feingold's hypotheses. *Can Psychiatr Assoc J.* 1978;23:241–248.
38. Harley JP, Matthews CG. The Feingold hypothesis: current studies. *J Med Soc N J.* 1979;76:127–129.
39. Wender EH. New evidence on food additives and hyperkinesis. A critical analysis. *Am J Dis Child.* 1980;134:1122–1125.
40. Mattes JA. The Feingold diet: a current reappraisal. *J Learn Disabil.* 1983;16:319–323.
41. National Institutes of Health. Defined Diets and Childhood Hyperactivity. NIH Consensus Statement. 1982;4:1–11.
42. Kavale KA, Forness SR. Hyperactivity and diet treatment: a meta-analysis of the Feingold hypothesis. *J Learn Disabil.* 1983;16: 324–330.
43. Werry JS, Sprague R. Methylphenidate in children: Effect of dosage. *Aust NZ J Psychiatry.* 1974;8:9.
44. 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;16:241–246.
45. Ward NI, Soulsbury KA, Zettel VH, Colquhoun ID, Bunday S, Barnes B. The influence of the chemical additive tartrazine on the zinc status of hyperactive children: A double-blind placebo-controlled study. *J Nutritional Med.* 1990;1:51–58.
46. Ward NI. Assessment of chemical factors in relation to child hyperactivity. *J Nutr Environmental Med*. 1997;7:333–342.
47. Huck SW, Cormier WH, Bounds WG, Jr. *Reading Statistics and Research*. New York: Harper & Row; 1974.
48. Friedman LM, Furberg CD, DeMets DL. *Fundamentals of Clinical Trials*. 3rd ed. New York: Springer; 1998.
49. Guy W. *ECDEU Assessment Manual for Psychopharmacology, Revised*. Rockville, MD: National Institutes of Mental Health, Psychopharmacology Research Branch; 1976.
50. Conners CK. Rating scales for use in drug studies with children. *Psychopharmacol Bull.* 1973;24:24–84.
51. Connors C. Connors' Rating Scales-Revised (CRS-R). In: *Handbook of Psychiatric Measures*. Task Force for the Handbook of Psychiatric Measures, ed. Washington, D.C.: American Psychiatric Association; 2000:329–332.
52. Gianarris WJ, Golden CJ, Greene L. The Conners' Parent Rating Scales: a critical review of the literature. *Clin Psychol Rev.* 2001;21: 1061–1093.
53. Sprague R, Christensen D, Werry J. Experimental psychology and stimulant drugs. In: Conners C, ed. *Clinical Use of Stimulant Drugs in Children*. Amsterdam: Excerpta Medica; 1974.
54. Etuk EM. *The Influence of Nutrition on Hyperkinesis* [dissertation]. San Diego, CA, United States International University; 1986.
55. Chernick E. *An Exploratory Study of the Effect of the Feingold Diet on Reading Achievement and Classroom Behavior* [dissertation]. Hempstead, NY, Hofstra University; 1980.
56. Breece AM. *Hyperactive Children: Changes in Hair Element Content*

with Diet Changes and Water Soluble Vitamin Supplements [dissertation]. Cincinnati, OH, The Union for Experimenting Colleges and Universities; 1980.
57. Alderson P, Green S, Higgins JP, eds. Cochrane Reviewers' Handbook 4.2.2 [Updated December 3, 2003] Sections 6.2 and 6.7. Available at: www.cochrane.org/resources/handbook/hbook.htm. Accessed April 8, 2004.
58. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials.* 1986;7:177–188.
59. Dunlap W, Cortina J, Vaslow J, Burke M. Meta-analysis of experiments with matched groups or repeated-measures designs. *Psychological Methods.* 1996;1:170–177.
60. Follman D, Elliott P, Suh I, Cutler J. Variance imputation for overviews of clinical trials with continuous response. *J Clin Epidemiol.* 1992;45:769–773.
61. *ES: A Computer Program for Effect Size Calculation* [computer program]. Version 1. St. Paul, MN: Assessment Systems Corporation; 1999.
62. Jennings AL, Davies AN, Higgins JP, Broadley K. Opioids for the palliation of breathlessness in terminal illness (Cochrane Review). *Cochrane Database Syst Rev.* 2001;4:CD002066.
63. Glass G, Stanley J. *Statistical Methods in Education and Psychology.* Englewood Cliffs, NJ: Prentice-Hall; 1970.
64. Rosenthal R. Parametric measures of effect size. In: Cooper H, Hedges LV, eds. *Handbook of Research Synthesis*. New York: Russell Sage; 1994:231–244.
65. *Stata Statistical Software* [computer program]. Version 6.0. College Station, TX: Stata Corporation; 1999.
66. Petitti DB. *Meta-Analysis, Decision Analysis, and Cost-Effectiveness: Methods for Quantitative Synthesis in Medicine*. New York: Oxford University Press; 2000.
67. Shadish WR, Haddock CK. Combining estimates of effect size. In: Cooper HM, Hedges LV, eds. *The Handbook of Research Synthesis*. New York, NY: Russell Sage Foundation; 1994:261–284.
68. Orwin RG. Effects of deficient reporting on meta-analysis: A conceptual framework and reanalysis. *Psychol Bull.* 1985;97:134–147.
69. Weiss B, Williams JH, Margen S, et al. Behavioral responses to artificial food colors. *Science.* 1980;207:1487–1488.
70. Mattes JA, Gittelman R. Effects of artificial food colorings in children with hyperactive symptoms. A critical review and results of a controlled study. *Arch Gen Psychiatry* 1981;38:714–718.
71. 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;26:56–61.
72. David TJ. Reactions to dietary tartrazine. *Arch Dis Child* 1987;62: 119–122.
73. Rowe KS. Synthetic food colourings and 'hyperactivity': a double-blind crossover study. *Aust Paediatr J.* 1988;24:143–147.
74. Pollock I, Warner JO. Effect of artificial food colours on childhood behaviour. *Arch Dis Child.* 1990;65:74–77.
75. Conners CK. A Third Challenge Trial. In: *Food Additives and Hyperactive Children*. New York: Plenum; 1980:77–85.
76. 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;125:691–698.
77. Goyette GH, Connors CK, Petti TA, Curtis LE. Effects of artificial colors on hyperkinetic children: a double-blind challenge study: "Experiment Two". *Psychopharmacol Bull.* 1978;14:39–40.
78. Goyette GH, Connors CK, Petti TA, Curtis LE. Effects of artificial colors on hyperkinetic children: a double-blind challenge study: "Experiment One". *Psychopharmacol Bull.* 1978;14:39–40.
79. Rose TL. *An Investigation of the Functional Relationship between Food Additives and Hyperactivity* [dissertation]. Gainesville, FL, University of Florida; 1977.
80. Swanson JM, Kinsbourne M. Food dyes impair performance of hyperactive children on a laboratory learning test. *Science.* 1980;207: 1485–1487.
81. Spring C, Vermeersch J, Blunden D, Sterling H. Case studies of effects of artificial food colors on hyperactivity. *J Special Education.* 1981; 15:361–372.
82. Harley JP, Matthews CG, Eichman P. Synthetic food colors and hyperactivity in children: a double-blind challenge experiment. *Pediatrics.* 1978;62:975–983.
83. Mattes J, Gittelman-Klein R. A crossover study of the effects of artificial food colorings in a hyperkinetic child. *Am J Psychiatry.* 1978;135:987–988.
84. Williams JI, Cram DM, Tausig FT, Webster E. Relative effects of drugs and diet on hyperactive behaviors: an experimental study. *Pediatrics.* 1978;61:811–817.
85. Sarantinos J, Rowe KS, Briggs DR. Synthetic food colouring and behavioural change in children with attention deficit disorder: a double-blind, placebo-controlled, repeated measures study. *Proceedings of the Nutrition Society of Australia.* 1990;15:233.
86. Levy F, Dumbrell S, Hobbes G, Ryan M, Wilton N, Woodhill JM. Hyperkinesis and diet: a double-blind crossover trial with a tartrazine challenge. *Med J Aust.* 1978;1:61–64.
87. Eich WF, Thim EB, Crowder JE. Effect of the Feingold Kaiser Permanente diet in minimal brain dysfunction. *J Med Assoc State Ala.* 1979;49:16–18, 20.
88. Conners CK. A Second Challenge Study. In: *Food Additives and Hyperactive Children*. New York: Plenum; 1980:55–68.
89. Conners CK. The Cookie Challenge. In: *Food Additives and Hyperactive Children*. New York: Plenum Press; 1980:45–54.
90. Thorley G. Data on the Conners Teacher Rating Scale from a British clinic population. *J Beh Assessment.* 1983;5:1–10.
91. Conners CK. Rating scales for use in drug studies with children. In: Guy W, ed. *ECDEU Assessment Manual for Psychopharmacology.* Revised ed. Rockville, MD: Department of Health, Education, and Welfare; 1976:303–312.
92. Conners CK. A teacher rating scale to use in drug studies with children. *Am J Psychiatry.* 1969;126:884.
93. Rapp DJ. Does diet affect hyperactivity? *J Learn Disabil*. 1978;11: 383–389.
94. Levy F, Hobbes G. Hyperkinesis and diet: a replication study. *Am J Psychiatry.* 1978;135:1559–1560.
95. Adams W. Lack of behavioral effects from Feingold diet violations. *Percept Mot Skills.* 1981;52:307–313.
96. Schachter HM, Pham B, King J, Langford S, Moher D. How efficacious and safe is short-acting methylphenidate for the treatment of attention-deficit disorder in children and adolescents? A meta-analysis. *Can Med Assoc J.* 2001;165:1475–1488.
97. Swanson JM, Kinsbourne M. Artificial Color and Hyperactive Behavior. In: Knights R, Bakker D, eds. *Treatment of Hyperactive and Learning Disordered Children: Current Research*. Baltimore: University Press; 1980:131–149.
98. Nain MP. Experimental toxicity of dye tartrazine. *Medecine et Nutrition.* 1983;19:365–374.
99. 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;156:557–561.
100. Corkum P, Rimer P, Schachar R. Parental knowledge of attention-deficit hyperactivity disorder and opinions of treatment options: impact on enrollment and adherence to a 12 month treatment trial. *Can J Psychiatry.* 1999;44:1043–1048.
101. McCracken JT. Attention-Deficit Disorders. In: Sadock BJ, Sadock VA, eds. *Kaplan & Sadock's Comprehensive Textbook of Psychiatry.* Vol 2. Philadelphia: Lippincott Williams & Wilkins; 2000: 2679–2688.
102. Loblay R, Swain AR. Adverse reactions to tartrazine. *Food Technology in Australia.* 1985;37:508–510.
103. Egger J, Stolla A, McEwen LM. Controlled trial of hyposensitisation in children with food-induced hyperkinetic syndrome. *Lancet.* 1992; 339:1150–1153.

104. Schettler T. Toxic threats to neurologic development of children. *Environ Health Perspect*. 2001;109:813–816.
105. Weiss B. Vulnerability of children and the developing brain to neurotoxic hazards. *Environ Health Perspect.* 2000;108:375–381.
106. Stein J, Schettler T, Wallinga D, Valenti M. In harm's way: Toxic threats to child development. *J Dev Behav Pediatr.* 2002;23(suppl): S13–S22.
107. Rose TL. The functional relationship between artificial food colors and hyperactivity. *J Appl Behav Anal.* 1978;11:439–446.
108. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 3rd ed. Washington, D.C.: American Psychiatric Association; 1980.
109. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 3rd, revised ed. Washington, D.C.: American Psychiatric Association; 1987.
110. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders.* 2nd ed. Washington, D.C.: American Psychatric Association; 1968.
111. Rowe KS, Rowe KJ. *The RBRI User's Manual: Parent- and Teacher-administered Inventories for the Assessment of Child Externalizing Behaviors, for use in educational and epidemiological Research.* Melbourne: Department of Pediatrics, University of Melbourne, Royal Children's Hospital, and Centre for Applied Educational Research, University of Melbourne; 1994.

Literary Quotes

Distractibility as an Asset

We tend to think of the temperament trait of distractibility as always being a liability because of its frequent association with problems in school performance. However, it can be an asset, especially in younger children, when it is more a matter of the ease with which the child's attention can be diverted from something unpleasant or inappropriate. It enables some children to be soothed or to soothe themselves more rapidly and effectively. Consider the poem "Portrait of a Boy" by Benét.

Steven Vincent Benét (1989–1943), noted American novelist, poet, and short-story writer, dealt mainly with American themes and fantasy. From high school many of us recall his "The Devil and Daniel Webster" or his Pulitzer Prize winning "John Brown's Body." Here, he invites us to share in the child's diverting fantasies as he deals with the pain of having been whipped.

After the whipping, he crawled into bed;
Accepting the harsh fact with no great weeping.
How funny uncle's hat had looked striped red!
He chuckled silently. The moon came, sweeping
A black frayed rag of tattered cloud before
In scorning; very pure and pale she seemed,
Flooding his bed with radiance. On the floor
Fat motes danced. He sobbed; closed his eyes and dreamed—

Warm sand flowed around him. Blurts of crimson light
Splashed the white grains like blood. Past the cave's mouth
Shone with a large fierce splendor, wildly bright,
The crooked constellations of the South;
Here the Cross swung; and there, affronting Mars,
The Centaur stormed aside a froth of stars.
Within, great casks like wattled alderman
Sighed of enormous feasts, and cloth of gold
Glowed on the walls like hot desire. Again,
Beside webbed purples from some galleon's hold,
A black chest bore the skull and bones in white
Above a scrawled "Gunpowder!" By the flames,
Decked out in crimson, gemmed with syenite,
Hailing their fellows by outrageous names
The pirates sat and diced. Their eyes were moons.
"Doubloons!" they said. The words crashed gold.
"Doubloons!"

In pediatrics we are learning more about the techniques of helping children to endure pain by self-regulation, using a variety of techniques such as self-hypnosis. From 65 years ago Benét seems to have given us a good literary example of the use of vivid mental imagery as a distraction to crowd out thoughts of a painful experience.

Benét SV. Portrait of a boy. In: Carhart GS, McGhee PA, eds. *Magic Casements.* New York, NY: Macmillan; 1937:323.

Submitted by William B. Carey, M.D.