Hyperactivity and Diet Treatment: A Meta-Analysis of the Feingold Hypothesis

Kenneth A. Kavale, PhD, and Steven R. Forness, EdD

This paper is a review of primary research investigating the Feingold hypothesis which suggests diet modification as an efficacious treatment for hyperactivity. The techniques of meta-analysis were used to integrate statistically the findings from 23 studies. The primary finding indicates that diet modification is not an effective intervention for hyperactivity as evidenced by the negligible treatment effects which are only slightly greater than those expected by chance. When the data were refined into groupings related to outcome and design variables, support was rendered for the primary finding. It is concluded that extant research has not validated the Feingold hypothesis and that diet modification should be questioned as an efficacious treatment for hyperactivity.

In 1975, Dr. Benjamin Feingold offered the hypothesis that the ingestion of artificial (synthetic) food additives (colors and flavors) and naturally occurring salicylates in foods results in hyperactivity and learning disabilities in children (Feingold 1975a, 1976). It was suggested that treatment be based on the Feingold Kaiser-Permanente (K–P) diet which is designed to eliminate all foods containing natural salicylates and artificial food additives from the diet (Feingold 1975b, Feingold & Feingold 1979).

Feingold (1974, 1975a, 1975b, 1976) reported that between 40% and 70% of hyperactive children who strictly adhered to the Feingold K–P diet demonstrated a marked reduction in hyperactive behavior. Although these conclusions were based solely on clinical observation and anecdotal accounts, Feingold (1975c), nevertheless, urged the immediate clinical application of the K–P diet to a Congressional Committee by stating, "It is not necessary to await the availability of basic data. It has been demonstrated that these children respond to dietary intervention" (p. 12).

Additional support was found in other uncontrolled clinical studies (Crook 1977, Keithly 1975, Stine 1976) which were characterized by positive results but serious methodological faults (Kolata 1978, Levine & Liden 1976, Spring & Sandoval 1976). A particularly scathing critique (Werry 1976) raised the ethical question of introducing a treatment before it has been shown to be efficacious; "I personally feel there is no greater breach of medical ethics than that of foisting a potentially worthless or dangerous treatment on a credulous public. Theirs may be the right to believe in magic and panaceas but ours as a profession is to act responsibly, cautiously, and scientifically, though not prejudicially" (p. 282).

Despite the subjective nature of Feingold's evidence and generally negative commentary by professionals, the hypothesis received widespread media attention and a favorable and enthusiastic response from the general public. In view of the lack of experimental evidence testing the Feingold hypothesis, two interdisciplinary groups were formed in 1975 to review the evidence for the presumed efficacy of the Feingold K–P diet in ameliorating hyperactive behavior. The Nutrition Foundation formed a National Advisory Committee on Hyperkinesis and Food Additives. This was followed by a second panel, the Interagency Collaborative Group on Hyperkinesis, convened by the U.S. Department of Health, Education, and Welfare.

The National Advisory Committee produced a report with the following summary statement; "The Committee concludes that data from critically designed and executed studies, free of the deficiencies of design noted, must be available before firm conclusions can be reached on the Feingold hypothesis" (Lipton 1975, p. 11). The Interagency Collaborative Group report (Kolobye 1976) discussed general guidelines and specific models for experimental research. After reviewing the available clinical evidence, the Interagency Collaborative Group concluded that studies to date "have neither proven nor disproven the hypothesis that a diet free of artificial colors and flavors reduces the symptoms in a significant number of children with the hyperkinetic syndrome" (p. 66).

With these reports suggesting that further investigation into the efficacy of the Feingold K–P diet was warranted, empirical studies were initiated but were marked by varying degrees of experimental rigor. In reviews of the available empirical evidence (Harley & Matthews 1980, Institute of Food Technologists' Expert Panel 1976, Wender & Lipton 1980, Stare, Whelan & Sheridan 1980, Tryphonas 1979, Wender 1977, Sheridan & Meister 1982), the conclusions have generally been ambiguous regarding positive findings for the Feingold K–P diet. No single empirical evaluation has reported consistent Feingold K–P diet effects for the reduction of hyperkinetic symptoms. Any reported positive findings have been either the result of *post hoc* analyses or no consistency between studies, both of which suggest the possibility of chance findings. Thus, the empirical evidence has not successfully clarified the effects of the Feingold K–P diet and has left unanswered the question posed by Stare, et al (1980); "Is a major dietary change justified on the basis of the modest behavioral improvement which may occur in a very small number of the children who met the somewhat vague criteria of hyperkinesis?" (p. 525).

The equivocal findings of the empirical literature have prevented rendering conclusive statements about the efficacy of the Feingold K–P diet. The reviews offered to date have been primarily narrative integrations resulting in impressionistic and subjective conclusions falling short of rigorous scientific standards for accumulating evidence. Jackson (1978) outlined the difficulties found in traditional research and suggested that conceptions of traditional research integra-

tion reveal it to be largely a matter of private judgment, individual creativity, and personal style. Such characteristics are inconsistent with the tenets of scientific research. Consequently, narrative reviews too often seem "like exercises in forcing an intransigent literature into the Procrustean bed of foregone conclusion" (Smith, Glass & Miller 1980, p. 36).

Glass (1976, 1977) proposed meta-analysis as a means of statistically integrating a body of literature and providing a rigorous alternative to the typical narrative discussions of research studies. The goal is to combine systematically the results of independent studies in order to obtain maximum information from existing research. Yet, "meta-analysis is aimed at generalization and practical simplicity. It aims to derive a useful generalization that does not do violence to a more useful contingent or interactive conclusion" (Glass, McGaw & Smith 1981, p. 23).

Meta-analysis is based on the *effect size* (*ES*) statistic which represents the quantification and standardization of treatment effects. It is defined by $ES = \bar{X}_T - \bar{X}_C / SD_C$ where $\bar{X}_T$ = mean of treatment group, $\bar{X}_C$ = mean of control group, and SD_C = standard deviation of control group.

Thus, treatment effects are transformed into a common metric (*ES*) independent of statistical significance. The resulting standardized mean difference, if positive, would favor the Feingold K–P diet, and, if negative, would indicate that the Feingold K–P diet is not effective. The *ES* statistic is comparable to a z-score and allows a similar interpretation if a normal distribution of responses to the Feingold K–P diet is assumed. An *ES* of +1.00 indicates that a treated subject at the 50th percentile would be expected to gain 34 percentile ranks and be better off than 84% of control subjects.

Meta-analysis has proven to be a valuable set of techniques for investigating problems in special education (Kavale & Glass 1981). The literature, both professional and popular, addressing the efficacy of the K–P diet, has intensified interest in the question. Because empirical evidence has proven less manifest than ethical and ideological positions regarding the effectiveness of the Feingold K–P diet, it becomes important to reassess the empirical literature through the application of research methods used to conceptualize, design, and analyze primary research. Consequently, the purpose of this study is to report a meta-analysis assessing the validity of the Feingold hypothesis. The primary question addressed: Is the K–P diet an efficacious treatment for hyperactivity?

METHOD

The initial step in a meta-analysis is to locate appropriate studies. Standard literature search procedures were used to identify relevant sources. This search yielded 25 studies but only 23 included a control group which was a necessary criterion for calculating the *ES* statistic. Of the 23 studies, 19 (83%) were from journals, while 4 (17%) were from books.

The next step was the quantitative description of study characteristics and their findings. First, to allow for the statistical description of relationships between the effect of diet modification and study attributes, data were recorded for substantive and methodological characteristics. Substantive features are characteristics specific to the problem studied while methodological characteristics are general concerns related to design and analysis. Next, study findings were quantified in the form of *ES* measurements. When means and standard deviations were reported, *ES* calculation was straightforward but, in the absence of these primary statistics, *ES* was obtained from solutions to t or F ratios, nonparametric tests, or percentages of improvement (via probit analysis) as outlined by Glass, McGaw and Smith (1981). Thus, these procedures allow for a full statistical integration of a collection of studies which includes not only a description of empirical findings on a common scale but also a description of how findings vary from study to study.

RESULTS

Although an attempt was made to locate all the studies in a domain, the studies located constituted only a portion of the population of studies. Ideally, the proportion of located studies is close to 100% but the obtained studies are neither a random nor probabilistic sample of the population. This situation results in the obtained sample presenting complex patterns of statistical dependence in the data set since a study may yield more than one comparison appropriate for calculating an *ES*. It is, therefore, necessary to decide upon the unit of analysis: the *ES* measurements themselves, irrespective of the number yielded by a study ("effect size"), or a single overall *ES* based on an average for all *ES* comparisons in the study ("study"). Particularly in a small sample of studies, the question arises as to the proper unit of analysis for aggregating findings. If "study" was the unit of analysis, then findings cannot be aggregated above the level at which many interesting relationships can be studied. With "effect size" as the unit of analysis, a problem surrounding independence arises; that is, how many independent units of analysis exist in a larger data set.

With the small number of studies in this meta-analysis, the decision regarding the proper unit of analysis is an empirical one; that is, do the findings differ when aggregated by "study" or "effect size"? The 23 studies yielded a total of 125 *ES* measurements; the average study produced 5.21 *ES* measurements with the range being 1 to 28.

When the unit of analysis is the study, the weighted average for individual studies *ES*'s ranged from −.954 to +.780, with the *ES* (average effect size) being .019. The standard deviation was .487 while the standard error (SE) was .102. (The standard error is a measure of the sampling instability of a mean and indicates how far the obtained mean might deviate from the population mean.) The median *ES* was .049 suggesting a skewed distribution with the *ES* probably underestimating the treatment effect. A 95% confidence interval, however, spans zero (−.180, .218), suggesting no treatment effect. Forty-five percent of the *ES* measurements were negative indicating an almost equal probability of control subjects showing an equal or greater treatment effect. If a normal distribution of responses to diet treatment is assumed (a convenient and unobjectionable assumption), the .019 standard deviation superiority for the average experimental subject indicated that diet modification moves the average subject from the 50th to the 51st percentile. This 1 percentile rank

gain suggests that the average experimental subject would be expected to be better off than 51% of control subjects, a level only slightly better than chance.

With "effect size" as the unit of analysis, the average across 125 *ES* measurements was .118 with a standard deviation of .415 and SE of .037. The range of *ES* measurements was −1.132 to +1.285 with a median of .045. Again, the distribution of *ES*'s is skewed with the mean probably overestimating the treatment effect. The similarity between "study" and "effect size" aggregation medians (.049 vs. .045) suggests that these values probably represent the "true" magnitude of treatment effects. A smaller proportion (27%) of the *ES* measurements were negative indicating that approximately three out of four *ES*'s revealed positive treatment effects. A 95% confidence interval (.045, .191) does not span zero indicating a reliable, albeit small, positive treatment effect. The $\overline{ES}$ of .118 indicates that Feingold K–P diet modification resulted in average benefits of slightly more than one-tenth standard deviation above control conditions. Figure 1 depicts this relationship as two normal curves separated by .118 standard deviations at their means. In relative terms, the .118 $\overline{ES}$ indicates that a child no better off than average (i.e., at the 50th percentile), would rise to the 55th percentile as a result of the Feingold K–P diet. Thus, the average subject was better off than 55% of control subjects at the end of treatment.

Although the two sets of findings are not dramatically different, which set, "study" or "effect size" is closer to the "true" picture? The means of "study" versus "effect size" aggregations (.019 vs. .118) were not significantly different ($t(146) = 1.01, p < .40$) indicating that the "effect size" mean (.118) was not larger than the "study" mean (.019). When *ES* measurements were correlated with the number of *ES*'s that the study produced, non-significant correlations were found for both "study" ($r = -.101$) and "effect size" ($r = -.078$) suggesting no relationship between the magnitude of effect and the number of comparisons in a study. Finally, if "study" and "effect size" are considered as classes, an intraclass correlation coefficient (p) indicates whether observations in the same class are related, or tend to be more like each other than observations in a different class. Thus, the larger the value of p, the more similar are observations in the same class, relative to observations in a different class. The value of p was .761 suggesting substantial relationship among *ES* measurements in the "study" and "effect size" classes. Although the value of p indicates "independent" data sets, Glass, McGaw and Smith (1981) showed how the intraclass correlation increases the variance error of the mean due to the nonindependence of findings within studies. The smaller standard error (.037) associated with the "effect size" integration suggests that the dependence in the data set did not produce an unsatisfactorily large standard error. Thus, any decrease in reliability of aggregates from what would be expected under independence is not sufficient to discontinue integrating data by "effect size."

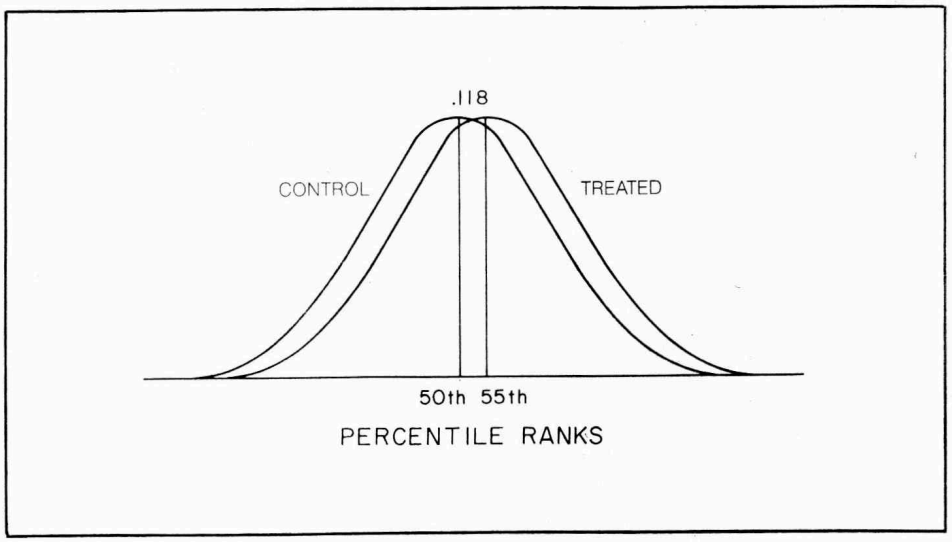

Figure 1. Effects of the Feingold Kaiser-Permanente Diet.

The $\overline{ES}$ of .118 was associated with a study that typically included subjects whose average age was 8.3 years and whose average IQ was 99.42. The average study was published in 1978 with a sample size of 38 (total N = 874) that was 87% male. Feingold K–P diet treatment lasted for an average of 39 weeks. No significant relationship was found between *ES* and number of subjects in a study ($r = .121$), objectivity of the hyperkinesis diagnosis ($r = .091$), or duration of treatment ($r = -.132$). A significant association ($r = -.255$, $p < .01$), however, was found between *ES* and age suggesting that larger *ES* measurements were found with younger subjects.

Although *ES* is readily comprehended by translation into notions of overlapping distributions of comparable percentiles, added meaning can be gleaned by reference to effects produced by other interventions. At a general level, in the elementary school, nine months of reading instruction would produce an $\overline{ES}$ of .67 while decreasing class size from 30 to 15 children would produce an $\overline{ES}$ of .15 (Glass & Smith 1979). Thus, the effects of a special intervention (Feingold K–P diet) are less than the effects of simply reducing the size of school classrooms. Perhaps a fairer comparison is with other special education interventions; psycholinguistic training was shown to result in a gain of .39 standard deviation units (Kavale 1981) while perceptual-motor training produced an $\overline{ES}$ of .082 (Kavale & Mattson 1983). The most revealing comparison would be with the most popular treatment for hyperactivity, stimulant drug treatment. Kavale (1982) reviewed 135 studies yielding 984 *ES* measurements. Only 15% of the *ES*'s were negative indicating an 85% positive response to drug treatment. The $\overline{ES}$ was .587 indicating that the average drug treated subject moves from the 50th to the 72nd percentile. This represents a 22 percentile gain compared to the 5 percentile rank gain for diet subjects. Drug treated subjects were better off than almost three-quarters of control subjects compared to just over one-half of diet treated subjects.

The average age (8.75) and average IQ

Table 1. Average Effect Sizes for Outcome Categories

Category	Number of Effect Sizes	Mean Effect Size	Standard Error	Percentile Equivalent
Conners scale - Parents	26	.156	.091	56
Conners scale - Teachers	9	.268	.138	61
Global improvement	23	.128	.108	55
Hyperkinesis rating	15	.293	.081	61
Attention	36	.015	.051	51
Disruptive behavior	6	.052	.138	52
Impulsivity	5	.153	.108	56
Learning ability	10	−.055	.140	48

(102) of drug treated subjects were similar to the averages for the Feingold K–P diet treated subjects, while the average duration of treatment was 18 weeks in the drug study compared to 39 weeks in the diet study. Thus, drug treatment is approximately five times as effective in less than half the time when compared to treatment with the Feingold K–P diet. These comparisons cast the Feingold K–P diet in an unfavorable light since it approximates the negligible effects of perceptual-motor training but is substantially lower than the positive effects of stimulant drug treatment.

The data were next aggregated into descriptive outcome categories with the findings shown in Table 1.

The effects of the Feingold K–P diet ranged from a loss of two percentile ranks (learning ability) to a gain of 11 percentile ranks (Conners Scale—teachers, and hyperkinesis ratings). A 95% confidence interval around the means for attention, disruptive behavior, and learning ability, span zero suggesting no treatment effect for these categories. Although teachers perceived more improvement than parents on the Conners Scale, the difference was not significant ($t(33) = .633$, $p < .25$) which is consistent with the finding of a high correlation between parent and teacher ratings (Zrull, Westman, Arthur & Rice 1966).

When the data from the Conners Scale were excluded, a significant difference ($F(5,89) = 2.66$, $p < .05$) was found among the remaining six outcome categories with Scheffe comparisons indicating that hyperkinesis ratings ($\overline{ES} = .293$) were significantly larger ($p < .05$) than the means for attention ($\overline{ES} = .015$) and learning ability ($\overline{ES} = -.055$). Thus, the Feingold K–P diet appears to have its only obvious effect upon overt behavior, specifically, a reduction in hyperactivity, and little influence upon more cognitive aspects of behavior.

Difficulties, however, exist with this conclusion. Caution must be exercised because of the problems with objectively operationalizing definitions of improvement on global ratings (Loney & Orduna 1975, Zrull, et al. 1966) and the psychometric deficiencies found with standardized rating scales (Barkley 1977, Sandoval 1977, Sulzbacher 1973, Werry 1978). These problems influence the "reactivity" or subjectivity of the outcome measure. Reactive instruments are those which are under the control of individuals who have an acknowledged interest in achieving predetermined outcomes. Nonreactive measures, on the other hand, are not easily influenced in any direction by the individuals involved. When *ES* was correlated with a rating of reactivity, there was a significant relationship ($r = .181$, $p < .05$) suggesting that larger treatment effects were associated with reactive measures. To examine the influence of reactivity, measures were grouped into two categories, reactive vs. non-reactive. Reactive ($n = 79$) and non-reactive ($n = 46$) measures exhibited $\overline{ES}$'s of .179 and .001 respectively which were significantly different ($t(123) = 2.37$, $p < .025$). Thus, these findings suggest that in those instances where instruments paralleled the valued outcomes of the experimenter, there was a tendency for those studies to reveal larger treatment effects. The point has been made that ecologic factors have often been ignored in determining educational outcomes (Forness 1981, in press).

The initial impetus for the Feingold K–P diet was based upon clinical observation and experience. The design is uncomplicated: Treatment A is given and effect B is measured. Although such quasi-experimental designs may provide impressive evidence, without the necessary experimental control the obtained evidence is merely suggestive and should not be perceived as compelling affirmative evidence for the hypothesis.

Experimental design requirements for studying psychotropic drugs have been outlined (Sprague 1978, Chassan 1967, Sprague & Werry 1971) and are applicable for evaluating the Feingold K–P diet. Of the 23 studies, six did not meet minimum design requirements and were essentially uncontrolled clinical trials. The six studies yielded 15 *ES* measurements with an $\overline{ES}$ of .337 and SE of .154 compared to the 110 *ES* measurements from the 17 controlled studies with an $\overline{ES}$ of .089 and SE of .037. Comparison of the $\overline{ES}$ found them significantly different ($t(123) = 2.16$, $p < .05$). The studies with no control exhibited a large positively skewed distribution of *ES* (median = .518) compared to a slight negative skew in the distribution of *ES* measurements from the controlled studies (median = .044). Additionally, there was a significant relationship ($r = -.193$, $p < .05$) between *ES* and ratings of design quality. The criteria for judging design quality (low, medium, high) were adopted from Campbell and Stanley's (1963) discussion of internal validity with sample size, method of subject assignment, extent of experimental mortality, and presence of measurement or statistical irregularities being the more important considerations. The significant correlation indicated that larger *ES*'s were associated with studies rated low on internal validity. These findings suggest that uncontrolled studies exhibit significantly greater treatment effects than controlled evaluations. Although psychologically persuasive, the findings from uncontrolled evaluations cannot be taken as evidence for the efficacy of diet intervention. The limitations inherent in uncontrolled studies, however, prevent them from proving the Feingold hypothesis. The lack of control makes it difficult to

attribute improvement to the treatment rather than to artifacts of the study conditions. Thus, a variety of alternative explanations, including the absence of diet monitoring, unclear sample descriptions, lack of control group, lack of "blind" procedure, inadequate consideration of placebo effects, and the failure to use reliable and valid outcome measures provide plausible reasons for the perceived treatment efficacy.

The controlled studies used two primary experimental methodologies: diet crossover studies and challenge studies. In a diet crossover study, groups of hyperactive children are placed on two different experimental diets. One diet follows the Feingold K–P diet recommendation while the other diet is disguised as the Feingold K–P diet but actually contains the substances supposedly eliminated. Because each group is tested on one diet, then switched to the other, and neither the subject nor experimenter knows the group assignment, this design is termed a "crossover" and "double blind" experiment. Since each subject receives each diet with the order of diet presentation randomized over time, double blind crossover studies represent a "within subjects" design. Subjects are evaluated while on each diet and then compared to assess the efficacy of the Feingold diet.

A challenge study selects a cohort of children who appear to respond to the Feingold K–P diet with reduced hyperactivity either in an earlier diet crossover study or clinical trial. This cohort is then divided into experimental and control groups. Both groups are given a strict Feingold K–P diet but the experimental group is also given a challenge food (usually a cookie or drink) that appears to meet the Feingold K–P diet guidelines but actually contains eliminated substances. The control group receives a control food that is similar in every way but does not contain the eliminated substances. The behavior of both groups is then evaluated and compared. In a challenge experiment, there are either two testing periods where experimental and control conditions are reversed so each subject serves as his own control or a single group is given both challenge and placebo food in a random sequence during a single testing period. The challenge studies include crossover and double blind elements but possess the advantage of ascribing behavioral change to the substances eliminated in the Feingold K–P diet.

Of the 17 controlled studies, seven used the diet crossover design and 10 were challenge studies. Aggregating *ES* data by experimental methodology found diet crossover studies yielding 45 *ES* measurements for an $\overline{ES}$ of .196 and SE of .063 while challenge studies revealed an $\overline{ES}$ of .045 and SE of .046 based upon 65 *ES* measurements. A 95% confidence interval around the $\overline{ES}$'s did not span zero for the diet crossover studies (.072 .320) but did include zero for the challenge studies ($-.046$, .136) indicating no treatment effect for the challenge study aggregation. A significant difference ($t(110) = 1.99$, $p < .05$) was found between the $\overline{ES}$'s for diet crossover and challenge studies.

Diet crossover studies, however, while an improvement over uncontrolled studies, possess methodological difficulties including: (a) inadequate control of placebo and social factors, (b) problems in keeping subjects and families from knowing the experimental diet (Feingold K–P) from the control diet, and (c) the possibility of carry-over effects for the experimental diet. Consequently, caution is necessary in interpreting the positive effects shown by the Feingold K–P diet on the basis of diet crossover studies alone.

A comparison of the median *ES*'s for diet crossover and challenge studies lends credence to this conclusion since the medians were similar (.048 and .037 respectively) and suggest a large negative skew for the *ES* distribution of diet crossover studies. Challenge studies, on the other hand, offer a methodology that permits the attribution of behavioral change to the substances eliminated in the Feingold K–P diet. Since challenge studies can be considered the "best" studies in terms of design and control, they offer the strongest evidence for the efficacy of the Feingold K–P diet. The treatment effect found in challenge studies produced a 2 percentile rank gain for the average experimental subject which is only slightly above a level of improvement expected by chance.

The aggregated findings for the two types of controlled evaluations combined with the significant relationship between *ES* and design quality suggest that the better the study the lower the *ES*. As a result of Feingold K–P diet modification, the subject who is no better off than average (i.e., at the 50th percentile) would be expected to rise to the 63rd percentile in uncontrolled studies compared to a rise to the 52nd percentile in the "best" controlled study (challenge design). The weight of the evidence suggests that the "truth" lies closer to the $\overline{ES}$(.045) found in the integration of findings from challenge studies. The "truth" indicates that at the end of Feingold K–P diet treatment the average experimental subject was better off than 52% of control subjects, a gain only slightly better than no treatment at all.

CONCLUSION

The findings of this meta-analysis do not offer support for the Feingold hypothesis. The evidence, when integrated statistically, indicated that the Feingold K–P diet produces a small treatment effect of approximately one-tenth of a standard deviation. This means that children placed in the Feingold K–P diet were better off than only 55% of control subjects at the end of treatment. Such a modest and limited gain suggests a more tempered view of the efficacy of the Feingold K–P diet for the treatment of hyperactivity than that asserted by Dr. Feingold and his proponents.

The refinement of *ES* measurements into more discrete groupings found that any appreciable improvement was related to overt aspects of behavior (i.e., symptoms of hyperactivity) while the more cognitive facets of behavior (e.g., attention, learning ability) revealed essentially no treatment effects. Consequently, a child placed on the Feingold K–P diet may exhibit slight improvement in behavioral functioning but not much else when compared to a child not treated with Feingold K–P diet. It is, therefore, important to examine the uncritical use of the Feingold K–P diet by parents to treat their children since it may postpone more appropriate medical, psychological, or educational intervention (Wender 1977).

Significant findings were related to the quality of research methodology found in the literature assessing the Feingold hy-

pothesis. In terms of research design, it was found that increased experimental rigor was associated with decreased treatment efficacy. Additionally, as the subjectivity of outcome measures increased, there was a parallel increase in perceived treatment efficacy. These findings offer support for the conclusion that food additives merely represent a placebo or a Hawthorne effect (Spring & Sandoval 1976, Swanson & Kinsbourne 1980a, Wender & Lipton 1980). If this assumption is correct, then it is possible that the lack of treatment efficacy found in well-controlled studies (i.e., those controlling for the placebo effect) may reflect the true state of affairs.

Although the Feingold K–P diet offers an appealing treatment approach for hyperactivity since it offers an alternative to stimulant medication and is consonant with attitudes labeling natural foods as "good" and artificial/synthetic ingredients as "bad," it is not without pragmatic difficulties specifically in terms of compliance. The Feingold K–P diet requires an abrupt change in lifestyle since increased vigilance is necessary in grocery shopping and food preparation, families generally cannot eat at restaurants, and the child cannot eat school lunches (Brenner 1977, Cook & Woodhill 1976, Sheridan & Meister 1982, Spring, Vermeersch, Blunden & Sterling 1981). Lew (1977) conducted a four week trial of the Feingold K–P diet on her family and concluded that "the Feingold Diet is indeed a very different and very difficult diet to maintain in practice. The deprivations to the participant(s) are real and is not the hyperactive child already set apart from his peers and family enough?" (p. 190). Additionally, the diet requires medical supervision since the restrictions may attenuate the intake of nutrients (particularly vitamin C) (Conners et al. 1976).

The methodology of meta-analysis, by applying the same objective methods used to analyze individual studies, provided evidence questioning the value of the Feingold K–P diet for the treatment of hyperactive children. By using scientific standards for accumulating evidence not found in less formal reviews, it was possible to draw reliable and reproducible conclusions suggesting that the Feingold K–P diet is not an effective intervention approach for hyperactive children. Yet, the widely publicized clinical evidence and quasi-religious belief espoused by "Feingold Associations" will make it difficult to depose the Feingold K–P diet as a treatment alternative for hyperactivity. Perhaps the empirical findings of this investigation will provide the basis for serious debate about the validity of the Feingold hypothesis and call into question the validity of the Feingold K–P diet as a treatment for hyperactivity.

REFERENCES

Barkley, R.A. A review of stimulant drug research with hyperactive children. Journal of Child Psychology and Psychiatry, 1977, 18, 137-155.

**Brenner, A. A study of the efficacy of the Feingold diet on hyperkinetic children: Some favorable personal observations. Clinical Pediatrics, 1977, 16, 652-656.*

Campbell, D.T. & Stanley, J.C. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally, 1963.

Chassan, J.B. Research Design in Clinical Psychology and Psychiatry. New York: Appleton-Century-Crofts, 1967.

**Conners, C.K. Artificial colors in the diet and disruptive behavior: Current status of research. In R. Knights & D. Bakker (Eds.), Treatment of Hyperactive and Learning Disordered Children: Current Research. Baltimore: University Park Press, 1980.*

**Conners, C.K. Food Additives and Hyperactive Children. New York: Plenum Press, 1980.*

**Conners, C.K., Goyette, C.H., Southwick, D.A., Lees, J.M. & Andrulonis, P.A. Food additives and hyperkinesis: A controlled double-blind experiment. Pediatrics, 1976, 58, 154-166.*

**Cook, P.S. & Woodhill, J.M. The Feingold dietary treatment of the hyperkinetic syndrome. The Medical Journal of Australia, 1976, 2, 85-90.*

Crook, W.G. Letter to the editor. News and Comment, American Academy of Pediatrics, January 12, 1977.

Feingold, B.F. Hyperkinesis and learning difficulties (H-LD) linked to the ingestion of artificial colors and flavors. Paper presented at the annual meetings of the American Medical Association, Chicago, June, 1974.

**Feingold, B.F. Hyperkinesis and learning disabilities linked to artificial food flavors and colors. American Journal of Nursing, 1975, 75, 797-803. (a)*

Feingold, B.F. Hyperkinesis and learning disabilities (H-LD) linked to the ingestion of artificial food colors and flavors. Paper presented to the U.S. Subcommittee on Health, September, 1975. (b)

Feingold, B.F. Why Is Your Child Hyperactive? New York: Random House, 1975. (c)

Feingold, B.F. Hyperkinesis and learning disabilities linked to the ingestion of artificial food colors and flavors. Journal of Learning Disabilities, 1976, 9, 551-559.

Feingold, B.F. & Feingold, H.S. The Feingold Cookbook for Hyperactive Children. New York: Random House, 1979.

Forness, S.R. Concepts of learning and behavior disorders: Implications for research and practice. Exceptional Children, 1981, 15, 497-509.

Forness, S.R. Diagnosing dyslexia: A note on the need for ecologic assessment. American Journal of Diseases of Children, in press.

Glass, G.V. Primary, secondary, and meta-analysis of research. Educational Researcher, 1976, 5, 3-8.

Glass, G.V. Integrating findings: The meta-analysis of research. In L. Shulman (Ed.), Review of Research in Education, 1977, 5, 351-379.

Glass, G.V., McGaw, B. & Smith, M.L. Meta-analysis in social research. Beverly Hills, Calif.: SAGE, 1981.

Glass, G.V. & Smith, M.L. Meta-analysis of research on the relationship of class-size and achievement. Educational Evaluation and Policy Analysis, 1979, 1, 2-16.

**Goyette, C.H., Conners, C.K., Petti, T.A. & Curtis, L.E. Effects of artificial colors on hyperkinetic children: A double-blind challenge study. Psychopharmacology Bulletin, 1978, 14, 39-40.*

Harley, J.P. & Matthews, C.G. Food additives and hyperactivity in children: Experimental investigations. In K. Knights & D. Bakker (Eds.), Treatment of Hyperactive and Learning Disordered Children: Current Research. Baltimore: University Park Press, 1980.

**Harley, J.P., Matthews, C.G. & Eichman, P. Synthetic food colors and hyperactivity in children: A double-blind challenge experiment. Pediatrics, 1978, 62, 975-983.*

**Harley, J.P., Ray, R.S., Tomasi, L., Eichman, P.L., Matthews, C.G., Chun, R., Cleeland, C.S. & Traisman, E. Hyperkinesis and food additives: Testing the Feingold hypothesis. Pediatrics, 1978, 61, 818-828.*

Institute of Food Technologists' Expert Panel on Food Safety and Nutrition. Diet and hyperactivity: Any connection? Nutrition Reviews, 1976, 34, 151-158.

Jackson, G.B. Methods for Reviewing and Integrating Research in the Social Sciences. Final report to the National Science Foundation for Grant No. DIS 76-20309. Washington, D.C.: Social Research Group, George Washington University, 1978.

Kavale, K.A. Functions of the Illinois Test of Psycholinguistic Abilities (ITPA): Are they trainable? Exceptional Children, 1981, 47, 496-510.

Kavale, K.A. The efficacy of stimulant drug treatment for hyperactivity: A meta-analysis. Journal of Learning Disabilities, 1982, 15, 280-289.

Kavale, K.A. & Glass, G.V. Meta-analysis and the integration of research in special education. Journal of Learning Disabilities, 1981, 14, 531-538.

Kavale, K.A. & Mattson, P.D. "One jumped off the balance beam": Meta-analysis of perceptual-motor training. Journal of Learning Disabilities, 1983, 16, 165-173.

Keithly, J. Report of pilot programs in schools in South San Francisco and San Mateo, California. Referenced in A.C. Kolobye (Chair) Interagency Collaborative Group on Hyperkinesis. First Report of the Preliminary Findings and Recommendations to the Assistant Secretary for Health. Washington, D.C.: U.S. Department of Health, Education and Welfare, 1975.

Kolata, B.G. Food additives and hyperactivity. Science, 1978, 199, 515-517.

Kolobye, A.C., (Chair). Interagency Collaborative Group on Hyperkinesis. First Report of the Preliminary Findings and Recommendations to the Assistant Secretary for Health. Washington, D.C.: U.S. Department of Health, Education and Welfare, 1976.

Levine, M.D. & Liden, C.B. Food for inefficient thought. Pediatrics, 1976, 58, 145-148.

*Levy, F., Dumbrell, S., Hobbes, G., Ryan, M., Wilton, N. & Woodhill, J.M. Hyperkinesis and diet: A double-blind crossover trial with a tartrazine challenge. The Medical Journal of Australia, 1978, 1, 61-64.

*Levy, F. & Hobbes, G. Hyperkinesis and diet: A replication study. American Journal of Psychiatry, 1978, 135, 1559-1560.

Lew, F. The Feingold diet, experienced (letter). Medical Journal of Australia, 1977, 1, 190.

Lipton, M., (Chair). National Advisory Committee on Hyperkinesis and Food Additives. Report to the Nutrition Foundation. New York: The Nutrition Foundation, 1975.

Loney, J. & Orduna, T.T. Using cerebral stimulants to treat minimal brain dysfunction. American Journal of Orthopsychiatry, 1975, 45, 564, 572.

*Mattes, J. & Gittelman-Klein, R. A crossover study of artificial food colorings in a hyperkinetic child. American Journal of Psychiatry, 1978, 135, 987-988.

*Mattes, J.A. & 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.

*Palmer, S., Rapaport, J.L. & Quinn, P.O. Food additives and hyperactivity: A comparison of food additives in the diets of normal and hyperactive boys. Clinical Pediatrics, 1975, 14, 956-959.

*Rapp, D.J. Does diet affect hyperactivity? Journal of Learning Disabilities, 1978, 11, 56-62.

*Salzman, L.K. Allergy testing, psychological assessment and dietary treatment of the hyperactive child syndrome. The Medical Journal of Australia, 1976, 2, 248-251.

Sandoval, J. The measurement of the hyperactive syndrome in children. Review of Educational Research, 1977, 47, 293-318.

Sheridan, J.J. & Meister, K.A. Food additives and hyperactivity. New York: American Council on Science and Health, 1982.

Smith, M.L., Glass, G.V. & Miller, T.I. The Benefits of Psychotherapy. Baltimore: The Johns Hopkins University Press, 1980.

Sprague, R.L. Principles of clinical trials and social, ethical and legal issues of drug use in children. In J. Werry (Ed.), Pediatric Psychopharmacology: The Use of Behavior Modifying Drugs in Children. New York: Brunner/Mazel, 1978.

Sprague, R.L. & Werry, J.S. Methodology of psychopharmacological studies with the retarded. In N. Ellis (Ed.), International Review of Research on Mental Retardation (Vol. 5). New York: Academic Press, 1971.

Spring, C. & Sandoval, J. Food additives and hyperkinesis: A critical evaluation of the evidence. Journal of Learning Disabilities, 1976, 9, 560-569.

*Spring, C., Vermeersch, J., Blunden, D. & Sterling, H. Case studies of effects of artificial food colors on hyperactivity. Journal of Special Education, 1981, 15, 361-372.

Stare, F.J., Whelan, E.M. & Sheridan, M. Diet and hyperactivity: Is there a relationship? Pediatrics, 1980, 66, 521-525.

Stine, J.J. Symptom alleviation in the hyperactive child by dietary modification: A report of two cases. American Journal of Orthopsychiatry, 1976, 46, 637-645.

Sulzbacher, S.I. Psychotropic medication with children: An evaluation of procedural biases in results of reported studies. Pediatrics, 1973, 51, 513-517.

*Swanson, J.M. & Kinsbourne, M. Artificial color and hyperactive behavior. In R. Knights & D. Bakker (Eds.), Treatment of Hyperactive and Learning Disordered Children: Current Research. Baltimore: University Park Press, 1980.

*Swanson, J.M. & Kinsbourne, M. Food dyes impair performance of hyperactive children on a laboratory learning test. Science, 1980, 207, 1485-1487.

*Trites, R.L., Tryphonas, H. & Ferguson, H.B. Diet treatment for hyperactive children with food allergies. In R. Knights & D. Bakker (Eds.), Treatment of Hyperactive and Learning Disordered Children: Current Research. Baltimore: University Park Press, 1980.

Tryphonas, H. Factors possibly implicated in hyperactivity: Feingold's hypothesis and hypersensitivity reactions. In R. Trites (Ed.), Hyperactivity in Children: Etiology, Measurement, and Treatment Implications. Baltimore: University Park Press, 1979.

*Tryphonas, H. & Trites, R. Food allergy in children with hyperactivity, learning disabilities and/or minimal brain dysfunction. Annals of Allergy, 1979, 42, 22-27.

*Weiss, B., Williams, J.H., Margen, S., Abrams, B., Caan, B., Citron, L.J., Cox, C., McKibbern, J., Ogar, D. & Schultz, S. Behavioral responses to artificial food colors. Science, 1980, 207, 1487-1489.

Wender, E.H. Food additives and hyperkinesis. American Journal of Diseases of Children, 1977, 131, 1204-1206.

Wender, E.H. & Lipton, M. (Co-Chairs), National Advisory Committee on Hyperkinesis and Food Additives. Final Report to the Nutrition Foundation. New York: The Nutrition Foundation, 1980.

Werry, J.S. Food additives and hyperactivity. Medical Journal of Australia, 1976, 2, 281-282.

Werry, J.S. Measures in pediatric psychopharmacology. In J. Werry (Ed.), Pediatric Psychopharmacology: The Use of Behavior Modifying Drugs in Children. New York: Brunner/Mazel, 1978.

*Williams, J.J., Cram, D.M., Tausig, F.T. & Webster, E. Relative effects of drugs and diet on hyperactive behaviors: An experimental study. Pediatrics, 1978, 61, 811-817.

Zrull, J., Westman, J. Arthur, B. & Rice, D. An evaluation of methodology used in the study of psychoactive drugs for children. Journal of the American Academy of Child Psychiatry, 1966, 5, 284-291.

*References used in the meta-analysis.

ABOUT THE AUTHORS

Kenneth A. Kavale *is an associate professor in the School of Education at the University of California at Riverside. He received his PhD degree in Special Education from the University of Minnesota.* **Steven R. Forness** *is a professor in residence at the UCLA Neuropsychiatric Institute. He received his EdD degree in Special Education from UCLA. Address: Dr. Kavale, School of Education, University of California, Riverside, CA 92521.*