Emanuela Taioli MD PhD
Director, Institute for Translational Epidemiology
Associate Director for Population Science, Tisch Cancer Institute
Director, Center for the Study of Thoracic Diseases Outcome
Professor, Population Health Science and Policy, and Thoracic Surgery

New York, 11/05/2020

Review of "Health Effects Assessment: Potential Neurobehavioral Effects of Synthetic Food Dyes in Children".

Based on my expertise and experience, I am reviewing the findings, assumptions, or conclusions I agreed I could review with confidence - point n 1 of Attachment 2, and to the extent possible - point n 3 of Attachment 2.

Conclusion n. 1 of Attachment 2: After reviewing the epidemiological literature on the neurobehavioral effects of synthetic food dyes, OEHHA concludes that the data suggest an effect of artificial food dyes on children's neurobehavior.

Review:

The contributors did an outstanding job in summarizing and interpreting the available epidemiologic data on neurobehavioral effects of synthetic food dyes. I have conducted a literature review myself, and have not found additional epidemiologic studies that needed to be included. I reviewed the comments on statistical analysis and the discussion of statistical concerns and found them satisfactory.

This is an exhaustive review of randomized clinical trials (RCTs) on dietary dye use and neurobehavioral effects in children. The choice of RTC was made *a priori*, and was based on a series of premises that I embrace. After reading and reviewing the available RCTs, I would consider adding a brief review of available observational studies, mostly because I realized that the RCTs were conducted on convenience sample of children, thus have limitations related to selection bias and external validity, and may not be representative of what happens in the general population. The other issue to consider is that the epidemiologic data are really scarce anyway, and perhaps a look at observational studies can improve the understanding and interpretation of the observed associations. I also would stress that data is really old, the most recent study included was conducted in 2007. These last 13 years may have introduced significant changes in dietary patterns; I'm not sure if the dyes currently used are exactly the same in quality and quantity as it was 10-15 years ago. This latter aspect should probably be checked out and verified.

The review includes studies both on general population of children as well as attention deficit disorder (ADD) children; both sets of studies show effects on neurocognitive function. Most studies test several dyes together, thus making it impossible to point at one of them as the responsible for the association with neurocognitive effects. This may

not be a negative thing all together, since it may represent an opportunity to look at the real life situation of a diet where many different components and dyes are ingested at the same time. This situation gives to opportunity to assess the overall effect that such diet has on cognitive behavior.

When the studies are described, Im not clear on why summary estimates are not presented, together with tests for heterogeneity. This would allow calculating an overall estimate, as well as identifying important sub-group and conducting sensitivity analyses. It is hard to judge an effect if there is no quantitative assessment of the data, and the reasoning behind the lack of a formal meta-analysis is not clearly articulated. If a scientific rationale is behind this choice, it should be described in a paragraph, otherwise an attempt of meta-analysis should be performed.

One thing that should be addressed clearly is the gap in knowledge, and this should be clearly stated in the conclusion. For example the lack of data on genetic susceptibility, the lack of data on biomarkers such as DNA methylation before/after exposure, the missing link between short term and long term effects of these repeated brief exposures on brain development and function are all important gaps that need to be highlighted and filled as soon as possible.

Another aspect that should be highlighted is that animal studies suggest some possible mechanisms, such as oxidative stress, or binding of the dyes to proteins that regulate neurotransmitters function. Given the complete lack of information on the metabolism, internal dose, and biological effects of these food components in humans, the information from animal studies become key to guide the development and the direction of future human studies. The hypotheses generated by animal experiments about possible mechanisms through which these dyes act on neurological functions should be listed as priorities to be investigated in order to better understand the risk for children and the possible remediation measures.

I understand from the Nigg et al meta-analysis, a well conducted study although a little old by now, that there are also concerns among the experts about the effects of exposure to food dyes alone versus dyes plus preservatives, with the latter being more harmful, but this point has not been addressed, unless I missed it.

There are two ongoing studies that have not generated results yet, but should be mentioned in future plans or somewhere appropriate. These data collections are ongoing and will eventually contribute relevant information that reflect more recent dietary habits and food composition:
A two arm randomized controlled trial comparing the short and long term effects of an elimination diet and a healthy diet in children with ADHD (TRACE study). Rationale, study design and methods.
BMC Psychiatry. 2020 May 27;20(1):262. doi: 10.1186/s12888-020-02576-2.

Rationale and design of an international randomized placebo-controlled trial of a 36-ingredient micronutrient supplement for children with ADHD and irritable mood: The

Micronutrients for ADHD in Youth (MADDY) study. Contemp Clin Trials Commun. 2019 Oct 26;16:100478. doi: 10.1016/j.conctc.2019.100478.

I have some more detailed comments that could help improve the document:

Page 30 – the authors indicate that no exclusion of studies was made based on the number of participants. Usually a minimum number of participants is set, for example 10, especially here where we are looking at randomized trials. The exclusion of very small studies could address possible publication bias, and could reduce heterogeneity. Study quality: the process for assessing study quality should be described more in detail. The table of items used to assess quality is not derived from one of the validated scores published by NIH, as far as I can tell. There are many validated systems for quality assessment, some for descriptive studies, some for RCT. I suggest that, unless there is a good reason for using a personalized quality score, the authors should use a validated published system. If the authors decide that they will use the current list of items for quality scoring purposes, then there should be a section that describes how the items composing the list were chosen, and how the list was validated.

Pag 32 – there is some confusion between the concepts of confounders and the quality scores. For example, there is a comment on elimination diets studies been possibly more sensitive in showing neurobehavioral effects than other studies; however this hypothesis was not confirmed by preliminary analyses, therefore the item was not included in the quality assessment. I think that these are part of sensitivity analyses, where subgroups are studied, and should not affect the quality evaluation of the studies.

Pag 38 – I cant figure out why elimination studies were excluded from the sub-analyses. Instead I suggest that they should be analyzed separately, as additional information may derive from these studies

Pag 39 – range of participants is quite large, and perhaps a decision to limit to studies with > 10 participants would have helped. Same for the dye dose, the range is quite large. Again, some strategy earlier when defining the inclusion criteria may have helped here, or the decision of conducting some sub-analyses of certain doses that are more meaningful and representative of the average dietary usage

Pag 40, 41 – several important concepts are included in this short section, and perhaps they should be separated into paragraphs with subtitles. We read here about dose response, latency and age groups. All these issues should be described more in detail, including implications. For example, is the dose response showing effects at doses that are commonly used, or only at doses that are unrealistically high? Is dose response present in certain age groups but not others? How about race? This section is a little bit of the core of the results and needs to be expanded and interpreted with more details. If there is no information on issues such as the ones I described above, then it should be stated as a gap in knowledge. I think that pointing at issues that haven't been studied is as important as showing results of studies that were properly conducted.

Discussion: the discussion is very thorough and detailed. I have no major comments.

Pag 44 – Design issues: I wonder how one can affirm with certainty that a RCT conducted on a convenience sample is superior to an observational study. This concept brings up the idea I discussed earlier that perhaps observational studies should have been given more weight, given the scarcity of available data.

Susceptibility: this is a very relevant paragraph, and again should be expanded. The first issue just touched upon is that younger children seem to be more sensitive than older children. Why is that? And would the increased susceptibility in younger children translate in any long-term effect on brain development deriving from this early sensitivity? If there is no scientific literature, then the gap should be highlighted. Genetic polymorphisms: looks like a metabolic chain is involved in degradation and elimination of these dyes, and clearly germline variations may play a role in individual sensitivity. The questions are: what is the degradation pathways of these dyes in humans? What are the genes involved? What is the population frequency of variants in these genes? I have read some of the relevant sections in this document (although were not among those assigned to me), just to have a better idea on what is known, and data seem scarce. Again, this is an important piece of the puzzle and if the data is missing, it should be mentioned.

Publication bias: Just as a suggestion, I wonder if these trials were registered in the public database? If so, there should be a way to find out how many of them were registered but not published.

The concluding statement about publication bias is not very convincing: the fact that several high quality studies show a positive association does not preclude the fact that other good studies with negative results were not published. Unless I'm not getting the point here, this statement is not completely appropriate and should be revised.

As I mentioned in the overall comment, there is no formal assessment of heterogeneity described here, and I wonder why. In general a test for heterogeneity helps defining the variability of the results, and points at subgroups and sensitivity analyses to try to address heterogeneity.

The conclusion should mention if there are long-term effects on these sensitive children, or on children in general who were exposed to these dyes. It seems relevant to talk about chronic, long term neurobehavioral effects. If there is no long term follow-up of these children, it should be mentioned as a scientific gap. I feel that transient effects could be of less relevance than persistent, chronic effects.

Tables: I would include a classic PRISMA graph to show included and excluded papers, together with number of papers excluded for each reason, and details about the reasons for exclusion. We usually group the excluded papers by broad categories, such as no RTC, etc

Table 2.2: I wonder if some of the data can be recovered from these papers, for example the first paper says that a fraction of subjects was challenged; can we recover these cases that underwent a challenge, and use the results for this review?

Table 2.3a: can sex be added as an extra column? It seems important to know if studies were conducted mostly in males, females, or both.

Table 2.4: looks like a sub analysis of US and UK is worth, since the majority of studies were conducted in these two countries. I wonder if a meta-analysis of dose-response can be conducted, given the number of studies reporting on it, from table 2.4

CONCLUSIONS

This is an impressive review of the topic of synthetic food dyes and neurobehavioral effects in children. Although there are no extra articles that I found for this review, I had suggestions on how to handle in a more formal way the various steps of the review, from defining the PRISMA for inclusion and exclusion of the studies, to quality evaluation, to summary estimates and sources of heterogeneity, to publication bias. All these aspects were discussed in detail above. I also suggest that gaps in current literature be highlighted so that future directions can be easily delineated to the reader.

Based on my expertise and experience, I am reviewing the findings, assumptions, or conclusions I agreed I could review to the extent possible - point n 3 of Attachment 2:

Conclusion 3 – Our estimates of exposure indicate widespread exposure to artificial food dyes in children, that children are exposed to a larger amount per body weight than women, and that the highest exposures were from over-the-counter medications in a single day.

Review:
I have some general comments on this section, for which I have some tangential expertise. This is a commendable effort to assess exposure levels in children, and characterize risks by poverty level, race and ethnicity, and education of the mother. Table 6.2 includes a review of available literature, and highlights how limited is the published information on human exposure. The table includes two unpublished thesis and five publications. This is a very small amount of data, with many inherent limitations for each of the studies reported. None of the studies seems to have reported age and sex-adjusted estimates, for example. I suggest that these limitations and gaps in literature are included in the comments.
Another big limitation of this section of the document is that all the assessment are estimates based on the dietary questionnaires, not actual measures, and derived from the known levels of various dyes contained in various food items as well as common drugs. There is no actual measure of blood levels, urinary metabolites, other markers of metabolism and excretion that could highlight levels of exposure and variations in such levels according to the important covariates mentioned above, such as age, sex, and race. This is a real problem, because all the variability observed and presented here is attributed to variations in dietary intake and in food dye content, not in individual ability to absorb, metabolize and excrete the products and their metabolites. The latter process, which is under control of several genetic pathways, could contribute greatly to the observed variability even if the dye intake is the same. I suggest adding this comment to the Summary section (6.9.2)
Another limitation that could be easily fixed is that the results of the new NHANES analysis, specifically commissioned by CALEPA, are reported in a very descriptive way, and is currently comprised of a series of univariate analyses. The available data very likely would allow for a more advanced statistical analysis, for example age, race and sex-adjusted estimates, at a minimum. The interaction between sex and age could also be looked at in detail. These further analyses can still be conducted and added to the document. I suggest doing so, as they would greatly improve a section that is very scarce in relevant information, mostly because there is very little literature available on exposure.

CONCLUSIONS
The section on human exposure to synthetic dyes in food should stress the limited number of studies available in the literature, and the many limitations of what is available. Among the most striking gaps in literature, we highlighted the lack of measures of dye levels in various human compartments (urine, blood), of metabolic gene polymorphisms that could contribute and explain individual variability in response

to exposure, the very descriptive nature of the statistical analysis of the NHANES data generated under CALEPA request. The section in my view should underline the existing gaps in knowledge, since human exposure assessment is one of the key steps in the evaluation process of any possible toxic substance.

DATE: November 9, 2020

PURPOSE: Scientific document review (in 10 pages)

SUBJECT: Health Effects Assessment: Potential Neurobehavioral Effects of Synthetic Food Dyes in Children Public Review Draft, August 2020. Hereinafter referred to as the "Report".

REVIEWED BY: This Report has been reviewed by Peter Spencer, PhD, FANA, FRCPath, Professor of Neurology and Occupational Health Sciences. Dr. Spencer is the former founding director of the Institute of Neurotoxicology at Albert Einstein College of Medicine, and of the Center for Research on Occupational and Environmental Toxicology and the Global Health Center of Oregon Health & Science University. He is a university-based, neuroscience-trained neurotoxicologist with decades of experience studying human and/or animal responses to exposure to chemicals/metabolites present in or added to food, and to exposure to various drugs, workplace and environmental chemicals with neurotoxic potential.

OEHHA TASK. In response to the Legislature's request, the California Environmental Protection Agency's Office of Environmental Health Hazard Assessment (OEHHA) conducted a multifaceted evaluation of the Food, Drug and Cosmetics Act (FD&C) "batch-certified" synthetic food dyes, focusing on seven of the nine food dyes approved by the US Food and Drug Administration for general use in food in the US. Batch-certification refers to a chemical analysis of each batch of dye sold to ensure that specific contaminants are below a legal limit. Color additives subject to batch certification are synthetic, derived from petroleum, and are listed on a product's ingredient label. The seven dyes are considered to contribute the greatest exposure to synthetic food dyes for the general US public. Specifically, the OEHHA was tasked with evaluating the literature designed to assess whether ingestion of food dyes affects the nervous system and behavior of children.

OEHHA REPORT. The study authors conducted a systematic literature search for and analysis of clinical trials examining the neurological effects of food dyes in children and preemptively assigned high confidence for conclusions drawn from the results of these studies. They also identified numerous experimental laboratory studies of mature and developing animals (rodents) designed to assess the effect of treatment with one or more synthetic food dyes. These studies included oral dye exposure during prenatal, infant, and juvenile development, with the examination of neurobehavioral effects in the offspring manifest during development and in adult animals. The effects of dyes were also evaluated with in-vitro high-throughput assay systems. OEHHA contracted with the University of California, Davis to estimate dye exposure from food and over-the-counter medications and vitamins intended for children. OEHHA also contracted with the University of California, Berkeley to combine these food dye levels with 2015–2016 NHANES data and to compute exposure estimates for a finer set of age groupings. Risk characterization compared these exposure estimates with US FDA Acceptable Daily Intakes (ADIs) derived during 1969–1987, and ADIs derived up until 2010 by the Joint FAO/WHO Expert Committee on Food Additives. Exposure risk was also characterized by poverty level, race and ethnicity, and education of the mother.

REVIEWER'S ANALYSIS. Data and conclusions derived therefrom were assessed for scientific validity. This was based on judgement of adherence to established principles of research practice as described in the many reports addressed in the Report. On occasion, individual studies were reviewed by examination of the original publications. Focus was placed on experimental design, study methodology, generated data, and conclusions drawn therefrom. Weaknesses in any of these areas were identified and used to assess the validity of the conclusions stated in the Report.

REVIEWER'S SUMMARIZED FINDINGS. OEHHA in concert with investigators at the University of California Davis and Berkeley carried out a comprehensive study of exposure/effect risk associated with oral intake by children of one or more of seven dyes used as coloring agents in food, medicines and vitamins. The OEHHA study focused on short and long-term risks to the human nervous system and behavior. Assessment was based largely on reports in the professional literature of relevant human, animal and in-vitro studies, the results of which were assessed in relation to contemporary human exposure estimates for Americans with different ethnic, racial, socioeconomic and educational profiles. The resulting August 2020 OEHHA Report for Public Comment represents a comprehensive approach that raises important questions about the safety of current

practices that expose children to the dyes under examination. The report lacks an assessment of the chemical structure of the seven dyes or their theoretically predicted and known metabolites. A more rigorous definition of chemical neurotoxicity is needed. The shortcomings in methodological design and data interpretation of some key human and animal studies should be pointed out, as should limitation in generalizability to the U.S. population. Consideration of neurobehavioral data in the context of data from other studies, notably those addressing the genotoxic properties of the food dyes/metabolites, is needed. Nevertheless, in general, this reviewer agrees with the broad conclusion that ingestion of food dyes may reversibly modify behavior in the short-term, which has special relevance to susceptible children in the context of Attention Deficit Hyperactivity Disorder. There is also scientific merit that certain of the seven synthetic dyes, notably the three azo dyes (Yellow 5, Yellow 6 and Red No. 40) and/or their metabolites, if genotoxic, may have potential to induce persistent nerve cell DNA damage and thus pose a risk for effects on brain function that appear later in life. Four synthetic dyes (Red No. 3, Red No. 40, Blue No. 1, and Green No. 3) appear to affect thyroid tissue, the function of which is required for childhood development, growth and neurobehavioral function.

REVIEWER'S SUMMARY ASSESSMENT OF REPORT CONCLUSIONS

The Reviewer supports the general conclusion" "*The scientific literature provides evidence in humans and animals, as well as mechanistic information, that synthetic food dyes may cause or exacerbate neurobehavioral problems in some children. Data from multiple evidence streams, including epidemiology, animal neurotoxicology, in vitro and high throughput assays providing mechanistic insight, taken together, provide support that some FD&C batch-certified synthetic food dyes impact neurobehavior in children. More evidence is currently available for Red No. 3, Red No. 40, and Yellow No. 5 than the other FD&C batch certified dyes*."

However, the conclusions for human studies rely to a significant extent on results obtained from U.K. and Australian studies of mostly Caucasian children, such that their relevance to the diverse population of U.S. children is unknown. Neurobehavioral changes were only identified in a proportion of children (so called "reactors"). Yellow No. 5 was specifically shown to trigger behavioral changes in so-called reactors. T**he conclusions for animal studies** rely on the aggregated results of diverse studies, several with weaknesses of methodological design, that together report changes in behavioral measures, neurotransmitter-related parameters, and other molecular systems required for normal brain development function. Evidence for induction of brain damage is not supported.

REVIEWER'S COMMENT ON "BIG PICTURE" QUESTIONS

(a) Absent an independent worldwide search of the literature, which is beyond the reviewer's assigned tasks, the Report captures the spectrum of neurobehavioral studies in children challenged with specific food dyes. Missing was presentation of the chemical structures of the dyes and the significance thereof, as discussed below.

(b) Absent an independent worldwide search of the literature, which is beyond the reviewer's assigned tasks, the Report captures available animal neurotoxicology studies relevant to the question of the neurobehavioral effects of the FD&C batch-certified synthetic food dyes. Several significant scientific issues relating to the methodological design of studies were not addressed, as detailed below. Two Iranian studies claiming the presence of neuroanatomical changes in the brain were overinterpreted as described below.

(c) Absent an independent worldwide search of the literature, which is beyond the reviewer's assigned tasks, the Report does not appear to have missed any studies that would inform a safe exposure level for neurobehavioral effects in children for any of the FD&C batch-certified synthetic food dyes. However, in regard to the three azo dyes, it would be valuable to examine their genotoxic potential because of emerging links between food exposure to natural toxins that form genotoxic compounds and induction neurodegenerative disease that may appear long after exposure has ceased.

1.BASIC PRINCIPLES

1.1. Chemical Neurotoxicity

From early development through adult life, there are many types of possible effects on the nervous system that result from overexposure to exogenous chemicals. While the descriptor "neurotoxicity" indicates adverse effects on the nervous system during development or throughout life, the term has shortcomings. The first

problem is that the descriptors “neurotoxin’ and “neurotoxicant” refer respectively to chemicals of natural or synthetic origin that have neurotoxic *potentia*l, but that potential is critically dependent on the subject (human, animal), the dosage and duration of exposure. Even substances required for normal physiological function, such as vitamin B6, have neurotoxic potential when the dosage is sufficiently large and prolonged. Conversely, exposure to minute amounts of high-potency chemicals (such as nerve agents) can be handled physiologically will no detectable adverse health effect. Other factors related to neurotoxic potential include species, sex, age, nutritional status, metabolic and metabolome status and, for humans, ethnicity and racial grouping. Ethanol, for example, has greater acute neurotoxic potential in persons with aldehyde dehydrogenase 2 deficiency, a genetic feature that is more common in Asian people than in other racial groups.

1.1.1. Developing Nervous System

The nervous system of the developing fetus is susceptible to chemical-induced changes as a function of the *stage of development* and secondarily the dosage. For example, substances that selectively interfere with molecular mechanisms required for cellular migration will express their neurotoxic potential *only* during periods of cellular migration. Importantly, chemical-induced changes in the developing brain have the potential for induction of *permanen*t adverse effects on neural structure and function.

1.1.2. Mature Nervous System

1.1.2.1. Pharmacological Effects. Single doses of neuroactive substances, such as the psychoactive substance caffeine, produce temporary and reversible changes on the nervous system function of adults. Both positive (brain stimulation, muscle strength) and negative behavior changes (sleep changes, anxiety) may result during the period of the chemical’s pharmacological activity. Mild physical dependence and withdrawal symptoms may accompany abstinence of caffeine intake. Substances with neuroactive potential, while not neurotoxic at low doses, are therefore relevant to an assessment of the health effects of food dyes. High doses of neuroactive agents or of agents that interfere with factors (such as glucose levels) critically required for normal neurological function can induce severe acute toxicity, including death. Additionally, chemicals that perturb the physiological balance of excitatory and inhibitory transmitters, or which act directly on their receptors, can induce acute and potentially chronic effects on neurological function. For example, the glutamate-receptor agonists beta-*N*-oxalylamino-L-alanine and domoic acid, which are present in certain plants and animals used by humans for food, can in certain doses induce neuroexcitatory effects that interfere with human motor and memory function, respectively.

1.1.2.2. Structural Damage. Single doses of some substances with neurotoxic potential may have delayed effects on the infant, juvenile and adult nervous system, with or without immediate effects. A classic example is the organophosphate compound tri-ortho-cresylphosphate which, with sufficient exposure, may have little or no immediate neurotoxic effect but, over the course of weeks, may induce distal degeneration of elongate nerve fibers in the spinal cord and peripheral nerves. Another example is carbon monoxide, a gas with acute toxic potential (narcosis, coma) but which may also precipitate a post-exposure and largely reversible movement disorder (e.g. Parkinsonism, dystonia, tremor) or even a progressive and fatal encephalopathy that appears weeks after the initial insult. Another example is the sugarcane mycotoxin 3-nitropropionic acid, which in large doses can induce coma and, upon reawakening, leave the subject with lifelong dystonia. These types of delayed-onset neurotoxic changes arise from structural damage to the nervous system.

1.1.2.3. Insidious Disease Onset. Repeated exposure to a large number of chemicals with neurotoxic potential may result in the insidious development of changes in nervous system structure and function; such changes are usually self-limiting after exposure ceases, and nervous system damage is repaired, albeit slowly by the re-growth of damaged structures. Examples include repeated oral exposure of rodents to the food flavor ingredient Musk Ambrette, which induces a peripheral neuropathy associated with distal nerve fiber (axonal) degeneration.

1.1.2.4. Long-latency Effects. There is growing evidence that single or multiple exposures to some chemicals may trigger molecular changes in the nervous system that do not surface clinically until years or decades later. This subject has particular relevance to progressive neurodegenerative diseases. One possible explanation is that exposure to culpable substances lowers the normal anatomical reserve of nerve cells but to a degree insufficient to surface in the form of clinical disease. However, with the addition of selected neuronal attrition

with the advance of age, certain damaged nerve cell populations eventually decline to a level that clinical disease surfaces. More recently, there is evidence that exposure to chemicals that induce certain types of DNA damage (i.e. genotoxins) may activate a silent pathological process that appears years or decades later in the form of a progressive neurodegenerative disease. The genotoxic property of food dyes is a subject investigated in relation to carcinogenic potential/risk but not for potential long-latency adverse effects on the nervous system. Nevertheless, identification of chemicals with genotoxic potential, whether or not they have been associated with experimental mutagenicity or carcinogenicity, has become relevant to safety assessment in relation to the nervous system.

2. ASSESSMENT OF NEUROTOXIC POTENTIAL

2.1. Study Design

Studies involving human subjects, animals or *in vitro* test systems should be based on a cogent hypothesis and effective design. In addition to the desirable controls discussed below, the study should have sufficient statistical power to provide a definitive answer to the hypothesis under study. A power analysis is used to estimate the minimal sample size based on a declared significance level, effect size, and statistical power. Studies that lack a sound study design may yield results with conclusions that are open to question.

2.1.1. Human Studies

A cardinal principle is adherence to a strong study design and observers who are blind to study interventions. This was not the case in the study by Bateman and colleagues (1987), which assessed the effects of artificial food coloring and benzoate preservatives on the behavior of 3-year-old children. Parents who were not blinded reported changes in behavior while validated psychological tests failed to register changes. A follow-up community-based, double-blinded, placebo-controlled food challenge reported replication of the first study (McCann et al., 2007), but the generalizability of these results in unknown. The two studies utilized children of families who lived in the U.K. Isle of Wight, which had a population of about 125,000 in 1991 of which, 2.7% were classified as “non-white” in 2011. Income levels were similar or somewhat lower than in other parts of the U.K. Whether or not the study groups were economically and ethnically representative of the population of the Isle of Wight, the results of these two studies are strictly only applicable to the subject population and have unknown relevance to other populations worldwide. For the U.S. population, the present Report concludes that: “Overall, non-Hispanic Black participants had significantly higher intake compared to other ethnic groups (Hispanic, non-Hispanic White, and Asian or other categories)” but that “Higher income was inversely, albeit weakly, associated with food dye exposure”.

2.1.2. Animal Studies

Species, strain, age, sex, nutritional status, route, method, dose, duration and purity of the administered article should all be controlled when designing and assessing the results of a study designed to measure the effects of a chemical on the nervous system. Most studies administer food and water *ad libitum*, but the composition of the diet and the presence of any contaminants in the diet, drinking water or administered article are rarely assessed. This has relevance because some food dyes contain contaminants with carcinogenic potential.

Studies are rarely performed with animals subjected to dietary restriction/excess, even though human subjects that eat food containing dyes are under/overfed. A minimum of three doses of the test article is required to assess dose-response but additional doses permit stronger assessment of a dose-dependent effect, a criterion commonly used as evidence that the test article was responsible for the outcome. Commonly omitted from experimental studies to assess the neurotoxic potential of one or more test articles are additional sets of animals that receive doses of a substance that is known (positive control) and is known not (negative control), to induce the neuro/behavioral effect of interest. Inclusion of a positive control compound that induces a response that matches previously published experience provides confidence both for the experimenter and for external observers that the test laboratory is performing to standard and, thus, the results from parallel studies of the test article(s) are credible. Further credibility is provided by the inclusion of a negative control compound that also performs according to previous experience. Since a negative control compound may itself induce changes in nervous system function, it is important to compare results generated by the test article not only with those associated with the test article’s vehicle (e.g. distilled water) but also with the effects of the negative control compound. A majority of studies seeking to assess the neurotoxic potential of test articles (including synthetic dyes) fail to include positive and negative control groups in experimental animal studies.

2.1.3. Comparison of Human and Animal Studies

Most studies utilized various mixtures of synthetic dyes to assess and define their effects on the nervous system and behavior. Such study designs yield results that can be very difficult, if not impossible, to interpret. Disparate chemicals and their metabolites may interact one with another in unpredictable ways, as illustrated by the following example. Peripheral neuropathy developed among adults who deliberately inhaled an organic solvent mixture for euphoric purposes. Since the solvent mixture was comprised mostly of methyl ethyl ketone (MEK), this was determined to be the culpable neurotoxic agent. However, animal studies demonstrated the culpable neurotoxic agent was a minor component of the solvent mixture (*n*-hexane) and that co-exposure to MEK potentiated the neurotoxic potency of *n*-hexane, while co-exposure to toluene did the reverse. This is an exceedingly rare example in which the interaction of concurrent exposure to two chemicals was delineated; in most cases, such interactions are rarely explored. and there are no examples of studies using mixtures of more than 2 chemicals that have yielded interpretable results. This example demonstrates the importance of controlled animal studies of individual substances in the interpretation of the effects of chemical mixtures on human subjects. Such studies are lacking for most of the seven synthetic dyes.

2.1.4. Cell/tissue Culture Studies

Studies of chemicals that test for effects in cell or tissue culture require special care in study design and interpretation. Direct application of a chemical to cells or neural tissues can elicit radically different responses from that seen when the same substance is administered systemically. Exposure *in vitro* is continuous. Whether the chemical applied to the in *vitro* system is metabolized over the course of exposure is rarely determined. The ability of the system to respond to chemical exposure in a manner that can be meaningfully interpreted is often questionable. While concentration-effect designs are commonly employed, positive and negative control compounds are often omitted. Exposure duration is often short and cellular responses may be non-specific. Determining whether there is any relationship between an *in vitro* observation and a behavioral effect in children is highly problematical.

2.2. Chemical Structure

The Report provides no information on the chemical structures of the synthetic dyes under review. Chemical structure is of cardinal importance because it may provide information on the presence or absence of a previously established active/inactive moiety for which effects may be forecast. Additionally, the structure of the chemical may provide information on probable metabolites and their potential for biological activity. An example of these principles is provided by the former food additive Musk Tetralin, the metabolite of which reacted with proteins to generate a blue pigment that predicted neurotoxicity in the form of nerve damage. Although the entire body of test animals turned blue after repeated treatment with Musk Tetralin, which indicated widespread reactivity with proteins, only the nervous system underwent pathological changes because of the unique architecture and functional requirements of neurons and their elongate axons. Subsequent studies of compounds related to Musk Tetralin demonstrated that the neurotoxic property was specifically dependent on the spacing of keto groups, such that 1,2-diacetylbenzene was chromogenic and neurotoxic, while 1,3-diacetyl benzene lacked both properties. This provides a clearcut demonstration that chemicals with closely related structures cannot be assumed to have comparable effects on the nervous system. As noted above, the simultaneous administration of multiple synthetic dyes adds a great deal of further complexity since the interactive effects the parent compounds and their metabolites cannot be predicted. Most studies analyzed in the Report tested chemical mixtures; only Yellow No. 5 (tartrazine) was tested as a single chemical.

Although the chemical structures of the food colors under review was omitted from the Report, mention is made of three so-called azo dyes, namely Yellow 5, Yellow 6 and Red No. 40, compounds that known to produce mutagenic metabolites. The first step in enzymatic metabolism of azo dyes appears to depend on the azoreductase activity of intestinal microbiota, the composition of which varies with ethnicity and other factors. Azo compounds/metabolites are of special concern because of their potential for genotoxicity that results in DNA damage and repair responses, oxidative stress, genetic instability, mutations, cell death and inflammation. Such properties have been previously related to cancer risks. Increasingly, there is concern that DNA damage/repair mechanisms may underly the genesis of certain sporadic neurodegenerative diseases that may incubate for years or decades prior to clinical expression. For this reason, it is important to consider the potential neurotoxic effects of azo dyes and their metabolites in relation to their genotoxic properties. While the

present report does not assess the genotoxic properties of azo dyes in relation to either carcinogenic or neurotoxic potential, there is sufficient information in the literature to express serious concern over the use of azo dyes food and medicine consumed by children. In accord with the precautionary principle as it relates to human health, it is recommended that azo dyes should not be used in food.

2.3. Chemical Access to the Nervous System

Chemicals that enter the blood stream have differential access to the nervous system as a function of chemical structure and developmental stage. In humans, the blood-brain regulatory interface, which is often described as a "blood-brain barrier" (BBB), matures within months of birth. Thus, compounds with neurotoxic potential have greater access to the nervous system of the developing fetus and newborn infant. However, even in the adult, certain substances can not only traverse the BBB but also enter brain tissue via the circumventricular organs (area postrema, median eminence of the hypothalamus, pineal gland, and the posterior pituitary), where there is normally no BBB in the infant, juvenile or adult human subject. Additionally, in the peripheral nervous system, a comparable blood-nerve barrier is normally absent in spinal and autonomic ganglia, such that chemicals circulating in the blood stream have immediate and direct access to both central and peripheral neural tissue.

Carotid artery injection of radiolabeled Red No. 3 in anesthetized rats resulted in radiolabel entering brain tissue (cerebral cortex, hippocampus, caudate, thalamus/hypothalamus) but the conclusion that chemical entry was solely via the BBB is likely incorrect. The hypothalamus is associated with specialized brain regions where the BBB-based capillary epithelium is fenestrated, such that free transfer occurs from blood to brain tissue. This possibility was recognized in a related study in which Red No. 3 was injected in the veis of conscious rats, when radioactivity was detected in 14 brain regions.

2.4. Secondary Effects on Nervous System Development

Chemicals that perturb thyroid function can interfere with normal brain development. Thyroid hormones are essential for brain development through specific time windows influencing neurogenesis, neuronal migration, neuronal and glial cell differentiation, myelination, and synaptogenesis. Red No. 3, Red No. 40, Blue No. 1, and Green No. 3 were active for an assay mapped to thyroid peroxidase (TPO), which measures TPO activity as a loss of signal. As stated in the Report, TPO inhibition could impair thyroid hormone synthesis, which ultimately could compromise neurodevelopmental processes. This raises a red flag for use of these food dyes in food materials consumed by pregnant women and infants.

2.5. Effects on Neural Receptor Function

The red and yellow dyes were all active in assays targeting dopaminergic and opioid receptor subtypes, with additional activity on G-protein-coupled receptors. Red No. 40 was also active for muscarinic and nicotinic cholinergic receptors. While these studies reveal the potential neuroactive properties of these substance, the significance of these observations in relation to the health of human subjects cannot be assessed. Few studies examined the effects of dyes on glutamate and GABA receptors, targets that would be associated with effects on the regulation of neuroexcitation and corresponding behaviors. Such studies require the use of positive and negative controls to aid in study interpretation. In general, however, transient effects of neuronal receptors are expected to be readily reversible.

3. FOOD DYE ASSESSMENT

3.1. Developmental Neurobehavioral Toxicology (DNT) Studies

DNT studies typically focus on detecting long-term or permanent effects on brain and brain function that occur after developmental exposure. Food dyes were provided at a fixed concentration in the diet throughout *in utero*, infant, juvenile and adolescent development and extending into adulthood. Studies performed in the 1970s-1980s yielded sparse evidence of adverse behavioral effects for Yellow No. 5, Red No.3, and Red No. 40. Later studies supported the hypothesis that sulfanilic acid, a common metabolite of the azo food dyes Yellow No. 5 and Yellow No. 6, was the effective agent in producing the dye mixture effects on activity. recorded in Japanese studies.

Animal studies performed in Japan investigated the neurobehavioral effects of several food colors, including Blue No 1, Yellow No. 5, Red No. 3, Yellow No. 6 and Red No. 40. These were dose-response studies, and

dosing was continued throughout several generations and life stages. While individual behavioral changes were noted, some related to sex, the Report notes several reservations that compromise the use of these studies for risk assessment.

Another research group administered a mixture of food dyes by gavage to rats during pregnancy and evaluated the offspring for behavioral effects at 90 days of postnatal age. Six of the seven FD&C dyes were represented (no Green No. 3). Three behavioral tests were used for adults. While the results of these three studies cannot be directly compared, they demonstrate long-term effects on behavior from *in utero* exposure at doses of the individual dyes found to have no effects in FDA regulatory reviews. Sensitive areas of brain function included regulation of activity, anxiety and exploration in a novel environment, and persistence in the forced swim test. Notably, no effects on learning and memory were seen. Analysis of brain receptors for neurotransmitters (glutamate, acetylcholine) were inconclusive.

Another study used a mixture of dyes (Red No. 40, Yellow No. 5, Yellow No. 6, Blue No. 1) that was added to drinking water of the male offspring after they had been weaned (PND 22) and throughout adolescence (PND 50). Measures of activity and anxiety were detected at earlier but not later animal ages.

In summary, these animal studies found changes in motor activity with mixed dye treatment but not in the results of learning and memory tests.

3.2. Adolescent/Adult Neurobehavioral Toxicity Studies

3.2.1. Brain Damage Underlying Neurobehavioral Toxicity

In the 1980s, two studies of neurobehavioral toxicity were conducted with dye exposures beginning at puberty or later. Two such studies used the azo dyes Yellow No. 5 and Red No. 40 to examine effects on cognitive function. Noorafshan et al. (2018) treated adult rats with/out Red No. 40. Treated animals showed more reference and working memory errors than controls, while learning the radial arm maze, and also in a retention test. Post-mortem examination of the brain was described as showing evidence of neuroanatomical changes (cell loss, dendritic shortening, reduced dendritic spines) in the medial prefrontal cortex that explained observed deficits in learning and memory tests of animals treated with high doses of Red No. 40. However, the reported neuroanatomical changes are not consistent with chemical-induced brain degeneration (*vide infra*). Similar reservations apply to the study of Rafati et al. (2017) who evaluated Yellow No. 5 using the same methodological design.

The claim that Noorafshan and colleagues induced structural brain damage in laboratory animals requires close scrutiny. In this study, adult Sprague Dawley rats received for 6 weeks daily gastric gavages of low doses (7 mg/kg/day) or high doses (70 mg/kg/day) of the azo dye Red No. 40 (purity unstated) dissolved in distilled water (the only control). Taurine was used as a protective compound. Behavioral tests began at 4 weeks and animals were anesthetized terminally at 6 weeks. Short and long-term memory deficits were found on tests of behavior of the high-dose group, and these were partially blocked by co-administration of taurine, an antioxidant and neuromodulator. Brains were fixed by cardiac perfusion (during life or after death?) with 4% (buffered?) paraformaldehyde. Volumetric and stereological studies performed on the fixed prefrontal cortex were reported to show reduced volume, loss of approximately half of the normal population of neurons and glial cells, a 40% reduction of dendritic length, and a 50-63% reduction of dendritic spines, in the high-dose group relative to the control group. Brain atrophy seen in high-dose animals was reportedly prevented by co-administration of taurine.

While this study was hypothesis-based and apparently carefully performed, the anatomical findings are not tenable and suggest study authors had limited neuropathological experience. Brain tissue does not undergo atrophy secondary to cell loss without a plethora of pathological changes that at any one time reflect a plethora of stages of cellular degeneration. Moreover, during a degenerative process, loss of neurons would be accompanied by an increase (not a decrease) in the number of glial cells, whether astrocytes, oligodendrocytes or microglial cells. The authors do not provide illustrations to support the claim of changes in dendritic length or dendritic spines in the brains of animals treated with high-dose Red No. 40 with and without taurine vs the vehicle-treated controls. In summary, the conclusion that neurocellular damage caused the observed behavioral changes is not supported, and evidence of dye-induced brain damage is lacking. In

contrast to the conclusion in the present Report, the reported brain histomorphology attributed to Red No. 40 and Yellow No. 5 does *not* help provide biological plausibility for the behavioral effects. As noted in the Report, other weaknesses of these studies relate to methods used to analyze data generated by the behavior studies.

3.2.2. Other Animal Studies

A rat study of Red No. 3 demonstrated a dose-dependent pattern of diminished activity that reached a low at 2 hours after dye administration and then returned to baseline by 7 hours. Noteworthy, the maximum effect on activity (2 h after administration) corresponded to the peak in Red No. 3 levels in circulation as described in a JECFA review. Post-mortem examination of brainstem, hypothalamus, hippocampus, and striatum revealed serotonin was lowered in a dose-dependent manner in all brain areas except the striatum. In the hippocampus, dose-dependent increases were shown for the serotonin metabolite 5-hydroxyindoleamine and the serotonin metabolizing enzyme monoamine oxidase A. (MAO-A). These results indicate that certain levels of Red No. 3 have reversible neuroactive properties in rats. As noted in the Report, the animal studies parallel challenge studies in children when behavior is measured shortly afterward a single dose of dye or mixture. In both rats and children, the effect of the dye peaks and then dissipates over a few hours post- exposure. However, with repeated exposure to rats, serotonin increased (rather than decreased) and MAO-A activity decreased (rather than increased) in all brain regions studied. These results are difficult to interpret.

Other studies that sought to examine the effects of various mixtures of multiple food dyes on animal behavior had variable experimental designs and results, such that conclusions drawn cannot be assessed with any degree of confidence. As discussed previously (*vide supra*), studies seeking to understand the effects of simultaneous exposure to multiple chemical substances suffer from fundamental problems in design and interpretation.

4. *In Vitro* Studies

4.1. Red No. 3 and Blue No. 1

Red No. 3 added to rat brain homogenates inhibited uptake of neurotransmitters including dopamine, serotonin, gamma amino butyric acid (GABA), and glutamate. Other *in vitro* studies showed that Red No. 3 can inhibit Na+/K+-ATPase, the enzyme activity of which is required to generate chemical energy to support the transport of neurotransmitters across plasma membranes. A proper balance of neurotransmitters is critical for maintenance of physiological homeostasis needed for normal brain function. Evidence that Red No. 3 is a potent inhibitor of glutamate uptake is significant because, in animal and human brains, accumulation of this excitatory neurotransmitter in extracellular space can trigger excitotoxicity leading to acute neuron damage, increased brain activity and, potentially, restlessness and even seizures. This potential neurotoxic phenomenon has not been studied in relation to increased motor activity in animal dye studies. Blue No. 1 showed no evidence of excitotoxicity when compared with glutamate as positive control in neuroblastoma cell cultures.

Whether and how results of *in vitro* studies relate to exposure *in viv*o will depend on the concentration of the free protein-unbound chemical in the bloodstream and its access to the nervous system. While there is a regulatory interface (BBB), between blood and brain tissue, there are several regions in the central and peripheral nervous system where chemical access to neural tissue is unimpeded (*vide supra*). This includes brain that regulate endocrine function, including thyroid hormone status, which appears to be impacted by treatment with Red. No. 3. As stated in the Report, “Studies are needed with oral administration and toxicocokinetics, including distribution and elimination of Red No. 3 and its deiodinated metabolites, di-iodo-fluorescein, mono-iodo- fluorescein, and fluorescein.”

Red No. 3 and Blue No. 1 respectively showed the highest and among the lowest overall activity in ToxCast assays, and Red No. 3 was the only dye with activity for monoamine oxidase. The ToxCast data support the estrogenic activity reported for Red No. 3. However, as noted in the Report: “While the ToxCast results did not provide overwhelming support for *in vivo* neurological alterations for the food dyes, data gaps and lack of biological coverage in ToxCast shine a light on areas to pursue.

4.2. Red No. 40

Red No. 40 was found to be the most potent of the azo dyes in an *in vitro* study specifically conducted for risk assessment of developmental neurotoxicity. The study was conducted in neuronal progenitor cells and looked at four of the seven FDA certified dyes, the three azo dyes (Yellow No. 5, Yellow No. 6 and Red No. 40) and the trimethylamine dye Blue No. 1. Red No. 40 stood out from the other FDA certified dyes because it reduced cell viability at micromolar concentrations.

4.3. Yellow No. 6

Yellow No. 6 reversibly inhibited the enzyme activities of human cholinesterase and pseudocholinesterase *in vitro.* Potency was about an order of magnitude lower than that of common organophosphate pesticides. Chemicals that inhibit the enzyme activity of acetylcholinesterase, which is required in physiological levels for normal neural and muscle function, are referred to as cholinesterase inhibitors. These are classified as reversible, irreversible, or pseudo-reversible. Reversible cholinesterase inhibitors are utilized for therapeutic purposes. In contrast, irreversible and pseudo-reversible inhibitors are often used in pesticides and biowarfare nerve agents.

5. Conclusion

Clinical trial studies provide limited evidence for short-term, reversible changes in behavior associated with exposure to food dyes in children. These are attributed to direct or indirect neuroactive effects.

Animal studies provide evidence for variable effects of exposure to food dyes on activity, memory and learning, sometimes with evidence of changes in brain neurotransmitter systems. The structural changes in the brain reported in two Iranian studies are not supported. While there is no evidence that exposure to food dyes at the doses tested elicits brain damage, given the genotoxic properties associated with the three azo dyes, studies are needed to assess the possibility of neuronal DNA damage resulting in genomic instability and long-latent effects require experimental evaluation.

Mechanistic and *in vitro* neurotoxicity studies of certain dyes provide evidence for the induction of oxidative stress, interaction with neurotransmitter receptors and key enzymes, and systems that exert influence on the brain including glucocorticoid pathways, thyroid and estrogen receptors. Of these, the effects of certain food colors on thyroid function, whether mediated by direct action on thyroid tissue or via the hypothalamus, has special importance for the developing human brain.

Data from multiple lines of evidence show that some FD&C batch-certified synthetic food dyes may temporarily impact neurobehavior in children. More evidence is currently available for Red No. 3, Red No. 40, and Yellow No. 5 relative to the other FD&C batch-certified dyes. The three azo dyes (Yellow No. 5, Yellow No. 6 and Red No. 40) are of special concern because of possible long-latency effects linked to genotoxic activity has not been ruled out. Brain and body development may be affected by food colors with potential to perturb thyroid function, and thereby brain development.

Future human and experimental studies should be hypothesis-based, employ an appropriately powered methodological design, and incorporate positive and negative control compounds to assist interpretation of effects observed with test articles.

6. Relevance to Attention Deficit Hyperactivity Disorder (ADHD)

Concerns about possible associations between synthetic food dyes and the exacerbation of ADHD symptoms in children prompted the Legislature to ask OEHHA to conduct this assessment. ADHD is characterized by symptoms of inattention, impulsivity and hyperactivity, and is considered to encompass a spectrum of neurobehavioral symptoms and severity. The percentage of US children and adolescents diagnosed with ADHD reportedly has increased from an estimated 6.1% to 10.2% in the past 20 years. It is postulated that environmental exposures to chemicals, including those used as food additives, may contribute to ADHD symptoms. While the results of key clinical studies of the effects of dyes on infant behavior have unknown relevance to diverse U.S. populations, there is a collection of animal studies that support the proposal that oral exposure to synthetic dyes can reversibly modify behavior in the short-term. The effects of oral exposure to dyes to genetic or chemical-induced animal models of ADHD have not been undertaken. Animal models

suggest of ADHD suggest involvement of dopaminergic, noradrenergic, and serotonergic systems, as well as more fundamental defects in neurotransmission.

Reviewer: Emily S. Barrett, PhD; Associate Professor of Biostatistics and Epidemiology at the Rutgers School of Public Health and the Environmental and Occupational Health Sciences Institute, Piscataway, NJ.

11/25/20

REVIEW OF "DRAFT HEALTH EFFECTS ASSESSMENT: POTENTIAL NEUROBEHAVIORAL EFFECTS OF SYNTHETIC FOOD DYES IN CHILDREN" (AUGUST 2020)

Based on my expertise and experience, I am reviewing the findings, assumptions, or conclusions I agreed I could review with confidence: Conclusion 1, and to the extent possible, Conclusion 3. Taken as a whole and to the best of my knowledge as an epidemiologist who studies chemical exposures and their impact on children's health and development, I believe this proposal to be based upon sound scientific knowledge, methods, and practices. Below, I highlight the factors that led me to that conclusion and identify particular strengths (and to a lesser extent, shortcomings) of the current report.

Conclusion 1

Conclusion 1 states: "After reviewing the epidemiological literature on the neurobehavioral effects of synthetic food dyes, OEHHA concludes that the data suggest an effect of artificial food dyes on children's neurobehavior."

The first section of the report consists of a systematic review of the scientific literature on seven synthetic food dyes that are approved for use by the US Food and Drug Administration (US-FDA) and are in commonly found in foods, beverages, over the counter medications, and vitamins: FD&C Blue No. 1, Blue No. 2, Green No. 3, Red No. 3, Red No. 40, Yellow No. 5, and Yellow No. 6. Based on these results, Conclusion 1 indicates that: (1) there is solid evidence that synthetic food dyes are associated with neurobehavioral measures (e.g. inattentiveness, hyperactivity, and restlessness) in children; and (2) some children may be particularly vulnerable to neurobehavioral outcomes following food dye consumption. However, it is also noted that the literature is variable, with associations observed in some studies but not others. These conclusions are consistent with the result of a 2012 meta-analysis on this topic as well (Nigg et al. 2012).

The process for undertaking the systematic review was overall sound and thoroughly described, though the inclusion of a "PECO" statement, a common element of systematic reviews, would have been useful. The reviewers used appropriate steps to identify publications of interest including searches in several of the largest biomedical literature databases as well as government reports. They justifiably chose to focus the systematic review on the results of clinical trials on this topic as they are considered the gold standard for strength of epidemiological evidence. Importantly, all of the studies reviewed employed a crossover design such that participants acted as their own controls, reducing potential confounding by relatively stable factors like socioeconomic

status. The strength of this design and applicability to this particular research question are explained well (for instance on p. 44) and the search strategy was well documented including the specific key words used in the search process, facilitating future replication (Section 1.3). Ultimately 27 studies were identified that met inclusion criteria for the review, and I am not aware of any additional studies that should have been included. The only lack of clarity noted in the inclusion criteria was #4 (p. 30) regarding "a neurobehavioral outcome related to hyperactivity or inattention was assessed"; a more comprehensive list of outcomes potentially "related to" hyperactivity and inattention would be preferable.

The quality of the 27 included studies was evaluated through a list of key factors to consider and a simple scoring system (Section 2.4). The list of included factors (2.4.1) is quite comprehensive, however it might have been useful to work within the framework of an existing Risk of Bias (RoB) tool intended for epidemiological studies. Such RoB tools (including, but not limited to, the Office of Health Assessment and Translation [OHAT] tool, Program on Reproductive Health and the Environment [PRHE]'s Navigation Guide, and the Integrated Risk Information System [IRIS] Tool) are specifically designed to assess internal validity by evaluating the extent to which elements of study design and conduct may have influenced results. While many such factors are captured in the scoring system devised by the authors, starting with and adapting an existing RoB tool might have added to the rigor of the systematic review. Alternatively, if extant RoB tools were considered but ultimately not used, it would have been helpful to explain the choice to instead create a new scoring system. Nevertheless, it is important to note that the factors used in study quality assessments (2.4.1) largely overlap with domains covered by RoB tools, thus the decision not to use an extant RoB tool is considered only a minor limitation and does not detract from the conclusions of the report. Overall Section 2.4 is an excellent summary of the decision making process around inclusion and exclusion of individual studies as well as the study elements that were then abstracted. In particular, I would like to note Section 2.4.3.9 in which the authors discuss consideration of magnitude of association as well as statistical significance as important evidence of causation. This is particularly important given the very small size of many of the studies considered, which may have been underpowered to detect effects. In fact, this reviewer questions the value of including extremely small studies (such as those with n=1), however this concern is ameliorated by the greater attention to and discussion of the larger and more rigorous studies.

The report is quite comprehensive in its data extraction and summaries. Tables 2.1-2.3 are helpful in ensuring transparency regarding excluded papers as well as data extraction and coding relevant to the 27 included papers. The overall approach utilized to select studies for inclusion and assessment of study quality was methodologically sound, however there was a lack of clarity on several minor points in Section 2.4.1, which explains the factors used to assess study quality. Clarifications needed include:

- On what basis was ≥50 mg/day used as a cutoff for a "high" dose?

- What constitutes an “adequate” washout period?

The Results section (2.6) is comprehensive and thoughtfully written, with consideration of a number of factors that might explain disparate results across studies including age of the study and the source of behavioral data (e.g. parent or teacher report, direct observation). However, it was somewhat surprising that differences in results across studies were not examined in relation to other factors, such as neurodevelopmental domain. While the studies focused on outcomes “related to attention”, some more granularity could be useful (for instance distinguishing between studies examining memory vs. activity). This was somewhat ameliorated by the recent Nigg et al. (2012) meta-analysis, in which neuropsychologists identified studies using tasks that specifically and directly measured attention; importantly, the effect size was stronger when including only those studies with that specific outcome.

Similarly, there was little consideration of whether results might vary based on the particular food dye used in the challenge, possibly because many studies used a mixture of several dyes making it hard to distinguish between their relative impacts. This omission may have been due to the paucity of studies examining a single, clear food dye exposure, as explained elsewhere in the report. Finally, the considerable differences in timing of exposure (as well as age at exposure) and latency until outcome measurement may contribute to inconsistent findings. Direct comparisons of studies with very similar designs (such as the Lok et al 2013 vs McCann 2007 comparison on pp. 43-44) are useful in parsing disparate results and could be employed more extensively in the report.

Despite these minor limitations, the Conclusion 1 remains well-supported, with the majority of studies reporting some evidence of association between food dye exposure and adverse neurobehavioral outcomes, despite differences in design elements, populations studied, and quality of research. Importantly, several of the more recent studies (which are among the highest quality and largest studies, including McCann et al 2007 and Bateman et al 2004) reported associations and went on to identify polymorphisms in histamine degradation genes that may underlie susceptibility to the adverse behavioral impacts of food dyes. The report appropriately highlights the results of these studies in multiple sections as they are among the most rigorous studies on the topic.

Several important elements of the current review that represent an advance beyond prior reviews (by the FDA and others) should be noted with regard to Conclusion 1. First, although prior evaluations focused particularly on the potential associations between food dyes and hyperactivity in children, in the current review, the committee also considers additional behavioral outcomes of interest. Second, recognizing that all children may be at risk, the committee evaluated studies in the general population as well as children with neurodevelopmental or behavioral disorders. Finally, although this external reviewer will not evaluate Conclusion 2, it is important to note that the

committee conducted an extensive review of the relevant animal toxicology literature that far exceeded prior reviews by the FDA.

In addition, this review points out several important limitations of the current epidemiological research in this area:

1) The majority of studies on this topic are quite old, which presents some issues. For instance, only two studies reported disclosures and source of funding, which is now common practice. There is potential for inherent conflicts of interest in industry funded research on this topic.
2) Similarly, a number of the studies were quite small. Of the 27 included in this analysis, 21 had samples sizes under 30 children, many of them less than 10 children. Although the report does a good job of considering both significant results and large effect sizes, there is a clear need for future work that is adequately powered.
3) There was considerable variation in the age of the children studied, and overall, there was some indication that effect sizes might be larger in younger children (e.g. preschool age) suggesting a need for additional study in this potentially vulnerable age group.
4) Most of the 27 included studies considered the combined effects of multiple food dyes, making it difficult to pinpoint which one or ones might be most strongly associated with behavioral issues. Additionally in some studies, another "agent" such as benzoic acid was used, potentially obscuring the true impact of the food dyes themselves (though importantly associations between food dye consumption and adverse behavior were reported in a number of studies that did not include such agents). Results of several studies of Yellow No. 5 alone (summarized in Table 7.10) suggest the need to conduct and compare studies of single food dyes to better identify those that might impact neurobehavioral outcomes.
5) There was a lack of blinding in many studies, which impact the child's own behavior as well as parental or researcher reports. Moving forward, direct observation by a psychologist who is blinded to the study arm (treatment vs placebo) would be the gold standard for outcomes measurement in this area.
6) Timing between exposure and outcome assessment was quite variable (and in some cases unclear) and there is a lack of clarity as to whether there may potentially chronic or long-lasting impacts of food dye exposure (particularly during sensitive developmental periods) on child neurobehavioral outcomes, as opposed to strictly adult impacts. While animal evidence suggests transient impacts, timing and type of exposure (acute vs chronic) clearly needs additional consideration in humans.
7)

I would also add, though it was not explicitly noted in the report, that given increasing evidence that chemical exposures may impact neurodevelopmental outcomes differently in males and females, sex differences in response to food dyes should be considered in future work. This hypothesis of potential sex differences in response to food dye exposure is further supported by some of the animal studies reviewed in Conclusion 2 (e.g. Tanaka et al 2001).

Regarding publication bias (discussed in 2.7.7), I concur with the reviewers that it is unlikely that publication bias would significantly skew the overall conclusions from this body of literature. While it is possible that some smaller studies with null or unexpected findings might not have been published, one would imagine that would be less of an issue with larger, well-designed trials. One possible exception would be the potential for large industry-sponsored trials showing associations between food dyes and problem behaviors being left unpublished. The addition of those studies, however, would only strengthen the overall body of evidence linking this exposure and outcome.

In summary, this reviewer affirms the quality of the systematic review of the epidemiologic literature, the results of which support Conclusion 1.

Conclusion 3

Conclusion 3 states: "Our estimates of exposure indicate widespread exposure to artificial food dyes in children, that children are exposed to larger amount per body weight than women, and that the highest exposures were from over-the-counter medications in a single day."

Conclusion 3, which evaluates children's level exposure to food dyes, is based on studies measuring food dye levels in foodstuffs, medications, and vitamins considered in concert with NHANES data on food consumption in children. It further examines exposure by demographic characteristics including poverty level, race/ethnicity, and maternal education. Based on the evidence presented, this reviewer concurs with the report's authors regarding Conclusion 3, namely that intake of synthetic food dyes is likely to be higher among children than adults and comes from disparate food sources including beverages, breakfast cereals, and desserts as well as from over-the-counter (OTC) medications and vitamins. Importantly, in novel analyses performed for this report, OTC medications were estimated to result in acute exposures that could exceed the FDA and Joint FAO/WHO Expert Committee on Food Additives (JEFCA) Acceptable Daily Intakes (ADI) even when used as recommended.

Conclusion 3 is supported by evidence from a variety of sources. Six recent studies have examined exposure to food dyes in U.S. and Canadian food stuffs either through: (a) dietary logs combined with ingredient lists or manufacturer information; or (2) direct chemical analysis of food items. Methods varied quite considerably across the six cited studies making it difficult to directly compare them, however in general, food dyes were commonly found in children's diets (or foods commonly consumed by children) and were particularly prevalent in certain food groups (e.g. fruit snacks, juices and soft drinks, candy). Of greatest relevance for estimating exposure in the general U.S. population are studies (e.g. Bastaki et al 2017) linking NHANES dietary data to estimated food dye content in those foodstuffs.

To complement and extend existing work, the authors chose to conduct an additional novel analysis for this report, which was well-justified and important for several reasons: (1) there are few population based studies on exposure to food dyes; (2) most exposure data are old and may not reflect current exposures among American children; (3) prior research didn't include additional potentially vulnerable populations like pregnant women; (4) prior research did not sufficiently consider additional sources of food dyes such as vitamins and medications. Novel chemical analyses conducted in a U.C. Davis laboratory in preparation for this report measured FD&C batch-certified food dye exposures in over the counter medications and vitamins. To my knowledge, this novel work is not yet peer-reviewed, and thus has not gone higher scrutiny by independent exposure scientists; nevertheless it is this reviewer's opinion that the new analyses greatly strengthen the overall conclusion due to the significant gaps in the prior literature.

In the new analysis, the researchers linked 2015-2016 NHANES demographic and 2-day dietary recall data (focusing on pregnant women, non-pregnant women of reproductive age, and children by age group) and food dye concentrations measured by the US FDA (Doell et al 2016, Harp et al 2013), to estimate food dye consumption (in mg/kg body weight/day) among NHANES participants using both typical-exposure and high-exposure scenarios. The estimates suggested the highest exposure occurred for FD&C Red No. 40 in children 9-16, 16-18, and pregnant women, with food dye consumption generally highest in children age 5-18, though for some dyes, like Blue No. 1 and Blue No. 2, estimates were highest for children ages 0-9). It should be noted that within each age group, only a fraction of NHANES participants actually consumed foods containing a particular food dye. For instance, among the 186 children under age 2, 108 (58%) consumed a food item containing Blue No. 2, while only 17 (9%) consumed a food item containing Green No. 3, thus for some groups and dyes, estimates were based on very small sample sizes. This was most notable for Green No. 3, which was consumed least frequently. The primary dietary sources of food dye exposure varied by dye and age group. For example, among the youngest children (0-<2), white icing was the predominant source of Blue No 1, whereas in older children, ice cream cones and soft drinks were more common sources. The food dye with highest exposure, Red No. 40, was most frequently consumed in fruit juice in children under 5 and in soft drinks in children 5-16.

In unadjusted analyses, total food dye consumption was weakly inversely correlated with higher income and income/poverty ratio and was highest in Non-Hispanic Black participants. Among adult women, food dye intake was higher in women with a high school degree (or GED) or less, compared to women with higher levels of education. While these results are interesting and may be a first step towards identifying populations that may typically have higher food dye exposures, I would consider these results preliminary and hypothesis-generating, rather than definitive given that no multivariable modeling was conducted. The discussion of these results in the report is tempered and appropriate.

As a next step toward risk characterization, the report compares FDA food dye intake under both typical-exposure and high-exposure scenarios (based again on NHANES dietary data) in relation to the US FDA and JECFA ADIs, with a Hazard index >1 indicating food dye exposure estimates (in mg/kg/day) exceeding the ADI (without contributions from medication or vitamins). Under both the typical-exposure and high-exposure scenarios, hazard ratios exceeded 1 for FD&C Red No. 3 among multiple age groups (children and pregnant women) and for both mean and 95% percentile exposure estimates (pp. 206-261). Estimates were typically highest for the youngest age group, children 0-<2 years. By contrast, hazard indexes were below 1 for the other food dyes under consideration.

With the addition of the novel food dye intake data from over the counter medications and vitamins (which had not been previously studied in this context), a second comparison to ADIs was made (p. 269). Notably, this set of comparisons did not include dietary intake of food dyes and thus would be an underestimate of typical total food dye intake. For certain brands of cold, cough, and allergy medicines intended for children, recommended use (based on the label) would result in Hazard indices for Red No. 40 or Blue No. 1 greater than 1 in children 6<12 and 12-16 (without any consideration of diet). Intake of other dyes in medication and through vitamins, by contrast, was estimated to be low. While use of these medications is likely to be intermittent for most children, there may be a subset who chronically use allergy medications with food dyes (potentially up to several times a day per instruction labels) and therefore may be particularly at risk of adverse neurobehavioral outcomes.

Finally the results of novel testing of food stuffs for food dye content at UC Davis further resulted in Hazard indices greater than 1 for some age groups based on consumption of a single serving of certain food items (or half a serving for children under age 2). Results were particularly notable for FD&C Red No. 3, for which a single serving of a variety of food items would result in a hazard index >1 based on the JECFA ADI (though not the US FDA ADI).

In conclusion, while the novel analyses of food dye intake through diet and medication use were not exhaustive in terms of the variety of foods and medications assayed, even with the limited scope of the new analyses, there is reason to believe that some children may routinely consume FD&C food dyes in amounts that exceed the US FDA and/or JEFCA ADIs, particularly through intake of OTC medications. Overall, this reviewer agrees with several of the noted limitations of the current literature on children's exposure and by extension, regulatory policy. Of particular importance are the observations that:

1) The older age of most of the studies reviewed (35-70 years old) is an important limitation of the literature, as there have been numerous advances in neurodevelopmental assessment since then, with more sensitive and rigorous tools now widely in use in the pediatric neurodevelopment literature.

2) The US FDA ADIs are estimated based on animal studies (on dogs and rodents) conducted in the 1960s-1980s, which are mostly not available for public review. To some extent, the WHO JECFA ADIs are based on more recent animal studies and the ADI for Red No. 3 in particular, was based on a study of adult human males and changes in thyroid hormone. However critically, for the WHO JECFA ADIs, as for the US FDA ADIs, none were based on neurobehavioral endpoints, making them inadequate for this purpose.

The report concludes, and this external reviewer agrees, that were the ADIs to be updated based on more recent data (where it exists) and on behavioral outcomes (rather than general toxicity), they would be considerably lower. This further suggests that current regulation of synthetic food dyes is out of date and not based on the most current evidence. Taken as a whole, I believe this proposal to be based on sound scientific knowledge, methods, and practices and have not identified any major weaknesses or omissions that would undermine the authors' conclusions.