Response to Peer Review and Public Comments on the August 2020 Public Review Draft "Health Effects Assessment: Potential Neurobehavioral Effects of Synthetic Food Dyes in Children"

Office of Environmental Health Hazard Assessment

April 2021

Response to Peer Review Comments

Responses to Comments from Peer Reviewer Emanuela Taioli, PhD, MD
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

"Comments on conclusion 1 of the request for review: 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."

Comment 1: "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."

Response 1: We thank the peer reviewer for her thoughtful review and comments. With regard to considering observational studies, the report describes 27 RCTs and evaluates these in detail. We do not think there would be added value to adding in an evaluation of the observational studies, given that the randomized blinded placebo-controlled clinical trials provide the more important data. The RCTs specifically evaluate the response to food dye challenges and avoid confounding as the subjects are compared to themselves, where the observational studies have more confounding and potential biases.

Comment 2: "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."

Response 2: We agree that on-going monitoring of food dye content and exposure is needed. Although the epidemiological studies are older, based on our research team's investigations, we have used the most comprehensive and high-quality food dye measurement data available in the U.S. for this assessment. The laboratory food dye measurements were overseen by US FDA laboratories on hundreds of food items selected based on their labeling of FD&C food dye ingredients (Harp et al., 2013). In addition, we used the most current NHANES food consumption data (2015-16) that was available for our exposure assessment.

To the best of our knowledge, the US FDA batch certification protocol has not substantively changed in the last decade, and the list of FD&C batch certified dyes are the same now as they were in 2007. The FDA has recently modified its use of X-ray fluorescence spectroscopy in order to report specific values for lead and arsenic for FD&C Yellow No 5 and its lakes; FD&C Yellow No 6 and its lakes; FD&C Red No 40 and its lakes; and FD&C Blue No 1 and its lakes (https://www.fda.gov/industry/color-certification/improvements-color-additive-certificates-analysis). We are not aware of any other changes in the quality of food color additives produced for the U.S. food supply in the last 10-15 years.

As discussed in Chapter 6 Exposure Assessment (Page 175), published reports have demonstrated that food dye production in the U.S. has been increasing since the mid-1950s, suggesting the possibility of higher population intake over time. However, the food dye measurement results (mg/kg food) derived from recent food sampling efforts performed by Dr. Alison Mitchell's laboratory at UC Davis as part of this assessment were in range with those reported by the US FDA laboratory data used in the Doell et al. 2016 assessment. The sampling and food dye measurements we performed in collaboration with UC Davis, however, were not broad enough to represent the US food pantry, and therefore we chose to base our current assessment on the most recent and high quality data that was available to us. Here is an excerpt from Chapter 6 Section 6.9.2 (Page 246):

> *"Overall, the new measurements of FD&C food dye concentrations reported by UC Davis for this assessment were within range or somewhat lower than the upper range of concentrations reported by FDA (Doell et al. 2016). However, UC Davis tested much fewer samples (~70 versus ~600) and their sample was intended to provide an independent check on current food dye concentrations reported in the scientific literature but not necessarily represent the full range of concentrations currently on store shelves."*

Comment 3: "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."

Response 3: We agree with the reviewer's first point, as already noted in our report, "Most studies involved administering multiple dyes at the same time so no single offending agent could be identified" (page 46). We also agree with the reviewer's second point, and have now added the following sentence to this section, "Regardless, studies involving mixtures will more closely represent real-life scenarios, where most children will be exposed to multiple dyes. In addition, while studies involving mixtures can make it more difficult to pinpoint any single causative agent, these studies can still be used to evaluate adverse impacts in the broad class of artificial dyes as a whole."

Comment 4: "When the studies are described, I'm 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."

Response 4: One new study was available since the Nigg et al. (2012) meta-analysis (page 43), and we provide an extensive review of this study (Section 2.7.4; page 43). This new study did not present means and standard deviations for analyses comparing placebo to artificial food dyes, and as such cannot be combined in meta-analysis with most other studies on this issue. This later point has now been added (page 43). As noted (page 43), we identified four other studies that were not in the Nigg et al. (2012) meta-analysis, but sample sizes were small and results were mixed. As such, their results are unlikely to cause a substantial change to the Nigg et al. (2012) results.

Comment 5: "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."

Response 5: The report includes a section on "Research needs and future directions", and this section mentions the need for further toxicokinetic information. We agree about the need for mechanistic studies, and studies on genetic susceptibility (e.g., following up on HNMT polymorphisms) and now further expand on this in the research needs section.

Comment 6: "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."

Response 6: As seen in Table 2.3a, only two studies tested dyes plus preservatives, while the very large majority did not. Many of the latter identified associations between these dyes and adverse effects on neurobehavior (e.g., Goyette et al., 1978; Levy et al., 1978; Rose, 1978; Rowe and Rowe, 1994; Pollock and Warner, 1990), providing evidence that the effects seen in the literature as a whole are not solely due to preservatives (page 46). This is now noted in our conclusion.

Comment 7: "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 reflects 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."

"Rationale and design of an international randomized placebo-controlled trial of a 36ingredient 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."

Response 7: Thank you for this information. We have not added this to the document as the results are not yet available.

Comment 8: "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."

Response 8: Table 2.5 shows that smaller studies were not more likely to report associations than larger studies. As such, it is unlikely that removing smaller studies would address any potential for publication bias. Studies with small sample sizes have had important impacts on public health in the past (e.g., thalidomide and birth defects, vitamin C and scurvy, diacetyl and bronchiolitis obliterans).

Comment 9: "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."

Response 9: The major advantage of the set of criteria we used is that they are specific for randomized clinical trials (RCT) of artificial food dyes and childhood neurobehavior. We developed our criteria by first starting with the National Toxicology Program's OHAT Risk of Bias Rating Tool (NTP Office of Health Assessment and Translation, 2019). We then improved this tool by making it more specific to the studies being reviewed. These modifications were based on basic epidemiologic principles (e.g., Rothman and Greenland, Modern Epidemiology, 1998), principles of good RTC design (Cummings, Hulley et al., Designing Clinical Research, 2013), a preliminary review of the relevant studies, issues raised in a public symposium sponsored by our agency, and concerns raised in critical review articles or reports. We have now added a more thorough description of the process we used to develop the set of quality criteria used in our report in section 2.4.1.

Comment 10: "Page 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.

"Page 38 – I can't 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."

Response 10: We did not remove all elimination studies, only those involving broad categories of foods. Although this was mentioned in our report, exclusion criterion number 2 has been rewritten for clarity. The passage now includes the following: "Studies that assessed the effects of a broad range of food groups and did not specifically evaluate synthetic food dyes. This includes elimination studies involving broad categories of foods. This was done because the eliminated foods could contain a number of different chemicals besides artificial food dyes. As such, any effect identified in these studies would be difficult to ascribe specifically to artificial food dyes."

With that said, we included two elimination studies that met our inclusion and exclusion criteria. However, the very large majority of studies we included were clinical dosing studies, not elimination studies. The elimination studies were removed from the analyses presented in Table 2.5 in order to reduce heterogeneity. However, the results and study quality criteria for these two studies are presented in Table 2.1, they are included in our overall description of studies (Table 2.4), and their findings are presented in our Results section (bottom of page 39).

Comment 11: "Page 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."

Response 11: Please see our response regarding sample size above. Sample size and dose levels were both included in the subgroup analyses we present in Table 2.5.

Comment 12: "Page 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."

Response 12: Most of the human studies did not present adequate information on dose response, and there was limited to no information on dose-response by age or race. Rowe and Rowe (1994) did describe a dose-dependent increase in response to the food dyes by children who were considered as "reactors". We describe this in chapter 7, Risk characterization. Detailed and consistent information on latency were also not presented in the studies.

As for dosing in the clinical trials, the exposures were given as total mg of mixture in the mixture studies, and not presented in mg/kg-day. Several study authors indicated they were attempting to mimic the exposures a child might experience from ingesting multiple dyes in the real world. To put the dosing in perspective relative to a typical ingestion of dyes, one can divide the mg of total dye in the mixture studies by typical body weights (from U.S.EPA Exposure Factors Handbook; available here: https://cfpub.epa.gov/ncea/risk/recordisplay.cfm?deid=236252). That can then be compared to total dye intake in our report, which is in Table 6.13. The mean of total dye intake across age groups (5 to 12 years of age) was 0.41 mg/kg-day and the 95th percentile was 1.3 mg/kg-day; the maximum was 11 mg/kg-day. We did not calculate total dye intake by age groups. Doses in the clinical trials we reviewed ranged from 15 to a high of 78 mg and were given to children of varying ages and body weights. In the McCann et al. (2007) study, the dose of Mix A administered to 3 yr olds was 20 mg. Dividing by the average body weight for 3 yr olds of 14 kg yields a dose of 1.4 mg/kg-day. For the 8 year olds, the dose of Mix A was about 0.8 mg/kg-d. These doses are in the range of what we estimated for total dye ingestion for children ages 5 to 16 in the U.S. As another example, Rowe and Rowe (1994) in evaluating the effects of Yellow No. 5 used doses of 1, 2, 5, 10, 20 and 50 mg in 54 children ages 2 to 14 years. These investigators found a dose-dependent effect on measures of activity in children. Using body weight by age from USEPA's Exposure Factors Handbook, those doses would then range from a low of about 0.02 mg/kg-d (lowest dose to oldest child) to a high of about 3.6 mg/kg-d (highest dose to youngest child). This can be compared to estimates of Yellow No. 5 exposure of about 0.056 up to about 0.3 mg/kg-day using the data in Appendix F, Table A6, on consumption of Yellow No. 5 by various age groupings of children, and dividing by reference body weights for those age ranges.

Comment 13: "Page 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."

Response 13: The RCTs specifically evaluate the response to food dye challenges and avoid confounding as the subjects are compared to themselves, where the observational studies have more confounding and potential biases. Though some studies found a stronger response in the younger children, this was not a consistent finding and we do not have adequate confidence in this finding to explore any etiologic reasoning further. Other than citing literature indicating long term behavioral implications of having attentional deficits during childhood, there is no literature that we are aware of specifically examining this issue relevant to food dye exposure. We have included these issues as data gaps in Section 8.6.

Comment 14: "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."

Response 14: There is very little information on the pharmacokinetics of the food dyes, and what is available is primarily from old animal studies. The azo dyes appear to be reduced in the GI by bacterial azoreductase and the resultant metabolites absorbed. At present, there is little information to inform the impact of genetic polymorphisms in the metabolic enzymes involved in metabolism of any of the dyes. In terms of response to food dyes, there is evidence of an association between histamine N-methyltransferase polymorphisms and response to the dyes (Stevenson et al., (2010). Histamine acts as a neurotransmitter in the brain adding to the relevance of this one study that found children with certain polymorphisms of histamine degradation to be more sensitive to neurobehavioral impacts from food dyes. The role of histamine in brain function including behavior is a new field of study and we have added some additional explanation of what is known into the document. We also note this issue as an area that should be explored in future research.

Comment 15: "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."

Response 15: We searched NIH RePort but did not find any studies specific for artificial foods dyes. We are not aware of any other public database where these studies might be available.

Comment 16: "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. "

Response 16: We agree that the statement is a bit confusing and not needed. We deleted that sentence. The previous sentence, "Given the widespread interest on the potential health effects of synthetic food additives, it seems somewhat unlikely that a number of well-conducted clinical trials would remain unpublished" is the primary argument.

Comment 17: "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."

Response 17: The Nigg et al. (2012) provides results for statistical tests of heterogeneity and Higgins I^2 values, and these are now mentioned in Section 2.7.1. However, given the considerable weaknesses of these tests, we were careful not to over-interpret them. In addition, the Chi-square test results presented in Table 2.5 of our report can also be used to evaluate heterogeneity within subgroups.

Comment 18: "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."

Response 18: To our knowledge, there are not long-term follow up studies of these children. However, since dye consumption is routine in U.S. children, then there will be repeated events of behavioral problems in those children who are reactors, which can negatively impact school performance, social interactions, and other aspects of the child's life. As noted in the executive summary, this may cascade into lifelong problems with social integration and success. Studies of children with ADHD have repeatedly shown long-term negative consequences of inattentiveness and hyperactivity into adulthood (Erskine et al., 2016; Klein et al., 2012).

Comment 19: "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."

Response 19: Although we did not generate a PRISMA chart, the inclusion/exclusion criteria for human studies are described in Chapter 2. As well, Table 2.1 concludes all the clinical trials we reviewed and Table 2.2 lists al the studies we excluded and the reasons for exclusion.

Comment 20: "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?"

Response 20: The data could not be recovered. This is now noted in Table 2.2. for Boris and Mandel (1994) : "Only a fraction of subjects were challenged with food dyes, and separate results were not provided for these subjects." As well, there were many other reasons that these studies were excluded, and it did not seem worthwhile to try to get data from these studies.

Comment 21: "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"

Response 21: Data on the numbers of boys and girls, when available, have now been added to this Table.

Comment 22: "Table 2.4: looks like a sub analysis of US and UK is worthwhile, 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."

Response 22: A subgroup analysis of studies done in the US vs. studies done elsewhere are presented in Table 2.5. Subgroup analyses based on dose level are also provided. Given that most studies involved different dyes or different combination of dyes, as well as differences in a number of major study design features (washout periods, dosing regimens, age groups, outcome metrics…), a meta-analysis of dose-response is not likely to be very informative.

Comments on 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.

Comment 23: "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

Response 23: On pages 180-186 of the report (Chapter 6), we present a detailed summary and comparison, including study limitations, of the two key published food dye exposure assessments that we examined, Doell et al. 2016 and Bastaki et al. 2017. As shown in Table 6.2, both of these studies stratified results by some age categories (Doell et al. 2016: US population: 2–5 years; Males 13–18 years old; General population ≥2 years; and Bastaki et al. 2017: US population aged: 2–5; years 6–12; years 13–18; years 19 and older).

As part of our discussion of these studies (Page 186), we describe how our expanded analysis compared to Doell et al. 2016 and Bastaki et al. 2017:

> "We estimated exposures for finer age groupings of children as well as pregnant women and women of childbearing age."

We agree with the reviewer that more research is needed to quantify sex differences in children's food dye consumption, and we have added this recommendation to Chapter 6 Section 6.9.2 (Page 246) of the report:

> "…Overall, on-going monitoring of food dye content in food is needed to determine exposure trends. Finally, future studies should evaluate differences in children's food dye exposure by sex."

Comment 24: "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)."

Response 24: Very little data are currently available. Future studies using biomarkers of food dye exposure (straights and lakes) to assess exposure are also needed.

As suggested by the reviewer, we have added these comments to the Chapter 6 Summary section 6.9.2 (Page 246).

> *"More research is needed on the pharmacokinetics (absorption, distribution, metabolism, and excretion) of ingested food dyes (straights and lakes) in children and adult populations. Currently, there is very limited research in this area. Future studies using biomarkers of food dye exposure are also needed to more accurately assess exposure. Overall, on-going monitoring of food dye content in food is needed to determine exposure trends. Finally, future studies should evaluate differences in children's food dye exposure by sex.*

Additionally, we have added the lack of data on exposure biomarkers, monitoring of food dye intakes, and evaluation of ingestion by sex into the discussion of data gaps in Section 8.6.

Our approach involved standard methods for estimating dietary intake. It was based on the NHANES food consumption data derived from standardized food frequency questionnaires whereby participants were asked to recall specific foods and the respective quantities consumed in the 24-hour period prior to an in-person interview. A second 24-hour recall dietary interview was then scheduled and conducted 3 to10 days later. An 8-digit U.S. Department of Agriculture (USDA) food code, linked to information about the type of product and brand, was assigned by NHANES to each reported food. This approach allowed us to quantify estimates of ingestion and is not related to pharmacokinetics.

We agree with the reviewer that more research is needed on the pharmacokinetics (absorption, distribution, metabolism and excretion) of ingested food dyes (straights and lakes) in children and adult populations. We state this in the section on Research Needs in chapter 8, section 8.6, which includes:

- Studies of the toxicokinetics of food dyes in humans, and studies of toxicokinetics of food dyes in animals using modern techniques and including exposures during *in utero*, preweaning, and juvenile stages;
- Studies examining absorption and bioavailability of straight versus lake food dye formulations are needed to inform the design and interpretation of exposure, toxicological, epidemiological, and clinical studies;

Comment 25: "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."

Response 25: For this report, we calculated children's food dye intake estimates stratified by five age categories (0-<2 years; 2-<5 years; 5-<9 years; 9-<16 years and 16-18 years). This was performed for seven the FD&C "batch-certified" synthetic food dyes. As shown in Tables 6.6 – 6.12, the "n's" (number of "eaters" per dye category, i.e., number of individuals that ate one or more foods containing a particular food dye) are presented by age category for Day 1, Day 2 and the 2-Day average. The "n's" for these age categories for the seven FD&C batch certified food color additives ranged from 9 (FD&C Green No. 3) to 860 (FD&C Red No. 40). The small n's in some of the categories makes it difficult to stratify by sex. While we agree that future assessments should consider sex-adjusted estimates, it would be beyond the scope of this report to also stratify our intake results by sex. As we stated above, we agree with the reviewer that more research is needed to quantify sex differences in children's food dye consumption, and we have added this recommendation to Chapter 6, Section 6.9.2 (Page 246) of the report:

> *"…Overall, on-going monitoring of food dye content in food is needed to determine exposure trends. Finally, future studies should evaluate differences in children's food dye exposure by sex."*

As part of our analyses for this report, we investigated socioeconomic differences in total food dye consumption in the US. Using women's and children's natural log-transformed total food

dye exposure estimates (mg/kg/day), we performed statistical analyses (Pearson correlations and T-tests) to investigate univariate associations between total exposure estimates and women's and children's poverty levels, race/ethnicity and women's level of education. Based on these analyses, we reported higher food dye intake among non-Hispanic Black participants compared to non-Hispanic whites. In Chapter 8.4 (Page 284), we state,

> *"Our analysis of socioeconomic determinants of food dye exposure suggest some weak trends with higher exposure in lower income families with less education, and significantly higher intake among non-Hispanic Black participants compared with other ethnic groups."*

Comment 26: "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."

Response 26: As noted in our responses above, we have added additional text stressing the limited nature of the data available for estimating consumption of synthetic food dyes.

Response to Comments from Peer Reviewer Dr. Peter Spencer, PhD, FANA FRCPath, Professor of Neurology and Occupational Health Sciences, Oregon Health & Science University
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.

Comment 1: "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."

"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."

Response 1: OEHHA appreciates the reviewer's recognition of the purpose and value of the OEHHA review and support for the conclusions of the review.

In terms of the definition of neurotoxicity, the reviewer states that a more rigorous definition of neurotoxicity is needed. In subsequent paragraphs, the peer reviewer walks through the various aspects of neurotoxicity. OEHHA agrees with the description of the various ways a chemical can impact the nervous system. In our search for information related to the food dyes, we captured all reported effects which included both functional and structural changes. This is broader than effects that can be directly associated with behavior, which we viewed as too narrow given the paucity of information about how an adverse effect on the nervous system may result in a behavioral change.

The reviewer also states that "the shortcomings in methodological design and data interpretation of some key human and animal studies should be pointed out…" The report does go through strengths and limitations of each of the animal and human studies described.

We address other specific points in the reviewer's summary quoted above in the following comments.

The reviewer also sets the review in the context of the broader, contemporary field of neurotoxicology and provides 3 "Big picture" comments (a,b, and c) to which responses are provided below. Additional comments from the reviewer related to these "big picture" comments are also incorporated below.

Comment 2: (page 3) (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."

(page 7) "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."

Response 2: Relevant information on chemical structures of dyes is interwoven with the text of Section 4 Toxicokinetics and Mechanistic Data. Section 4 mentions the main chemical classes of dyes we reviewed (azo, triphenylmethane, and xanthene dyes) common structural and functional properties involved in their absorption metabolism and distribution and potential mechanisms of action. Section 4 discusses common structural properties including multiple aromatic rings that produce colors in the visible spectrum, water solubility often achieved by sulfonation of the dyes that allows distribution in food matrices, and the wide-spread protein binding which allows the dye to produce uniform, permanent color. In particular, basic research with contemporary methodologies is currently revealing thermodynamic properties of dye binding with specific proteins of interest and resulting alterations in their functioning. These papers are cited but not discussed in detail. We have added a section on the chemistry of the dyes, including a figure with the chemical structures, to the Introduction.

The dependence of toxicity on even small molecular changes also emerged from our literature review. Current development of drugs based on dyes that target purinergic receptors and prevent beta-amyloid aggregation is demonstrating the biological impact of small modifications of dye structure.

We agree that our analysis of the available studies is limited in our ability to fully address additive or interactive effects of dyes. However, nearly all of the epidemiologic studies and many animal studies, that were designed to mirror human studies, were mixed dye administration studies reflecting real life exposure scenarios.

Comment 3: (page 3) "(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".

(page 10) "**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."

"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."

Response 3: The reviewer is correct that the articles (Noorafshan et al. 2018, Rafati et al., 2017) do not provide data supporting a neuropathological interpretation of chemically induced brain damage. However, the authors' interpretation is appropriately limited to "structural change".

Unfortunately, the study did not include a histopathological examination of the brain region (medial prefrontal cortex). Important qualitative information needed for clinical diagnostics is not provided by the histomorphometric stereology. The neuropathologist has extensive knowledge of tissue architecture and histology, response to injury, and pathophysiology, none of which is incorporated in histomorphometry measures. As stated by practitioners "Histopathology remains the gold standard in tissue evaluation and pathologists should embrace these quantitative techniques, rather than viewing them as a possible replacement" (Brown 2017). As an example, neuronal number is derived in unbiased stereology by multiplying the count density in the sampled sections by the total volume of the mPFC. Regional changes in the architecture and composition cannot be taken into account. Localized cell loss and accompanying compensatory responses that might identify neurodegeneration would not emerge. Thus a neuropathological conclusion does not emerge. However, the histomorphometric stereology data is valid and can be appropriately characterized as "structural change" awaiting actual neuropathology studies. Thus, we continue to believe that these studies provide support for biologic action of the synthetic food dyes on brain tissue and provide biological plausibility for the behavioral effects noted in the studies.

Comment 4: (page 4) "(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."

(page 5) "**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."

(page 6) "**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."

(page 8) "… 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."

Response 4: The reviewer here introduces a possible mechanism of action for food dye neurotoxicity which did not emerge from the OEHHA literature review. This mechanism is based on the idea that genotoxicity in non-proliferating tissues like nerve tissues at developmental life stages would have a distinctive profile of late onset neurodegenerative disease (Kisby and Spencer 2011).

None of the studies reviewed by OEHHA would have detected a neurodegenerative disease that appears after a considerable lapse in time following a developmental exposure. There is a general lack of research on long term neurobehavioral effects of developmental exposure to food dyes. Only three studies we reviewed used a developmental exposure discontinued before maturity and neurobehavioral evaluations in adults (Doguc et al. 2015, Doguc et al. 2019, Erickson et al. 2014). Also, none of the developmental neurotoxicty studies evaluated genotoxicity at the time of exposure.

As the reviewer states, the FDA requires data on genotoxicity and carcinogenicity for proposed synthetic food dyes. There is no data requirement for neurotoxicity or developmental neurotoxicity guideline studies. However, neither of these guideline study protocols would identify "long-latency adverse effects" on the nervous system.

The FDA certification protocol does not require periodic reviews of the literature to identify new toxicology research. However, WHO/JECFA does review the literature at intervals and concern has arisen from new information on genotoxicity of food dyes, leading to request for new studies and revision of ADIs. OEHHA did acquire 69 research reports on genotoxicity of food dyes through its literature review. Recently, concern over genotoxicity of dyes has led to in vitro genotoxicity testing using contemporary methods like Comet assay and micronuclei assay, and focusing on the azo dyes Red No. 40, Yellow No. 5 and Yellow No. 6 . OEHHA did not find any link between genotoxicity and neurotoxicology during our literature review. The only avenue for using genotoxicity as a mechanism of food dye neurotoxicity in informing public health policy would be via investigator-initiated research.

We don't have detailed information on what contaminants are present in individual batches of dyes and whether they have been evaluated for genotoxicity. FDA has guidelines for screening for genotoxicity of "impurities" in drugs using a combination of structure-activity analysis and in vitro testing (Kruhlak et al. 2017) but this has not been applied to food dyes. FDA research at NCTR (Chung and Cerniglia 2007; Feng et al. 2012) has explored genotoxic structure activity relationships of dyes and their metabolites to identify chemical groups involved that might also be present in contaminants. Contaminants can include raw materials used to synthesize the dyes, by-products of the synthetic reactions, breakdown products and accidental contamination. Some of these chemicals are known to be genotoxic but many have not been studied.

Comment 5: (page 6) "**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."

Response 5: We agree that it is possible that effects vary by race, ethnicity, or metrics of socioeconomic status. However, adequate data to thoroughly address these issues are not available. We have now noted this in our section on research needs. While the Isle of Wright studies may not be applicable to every child worldwide, we are not aware of any evidence or reason to believe they are not applicable to similar groups of children outside this locale. This is supported by the fact that effects similar to those seen in the Isle of Wright studies were also seen in studies done in other parts of the world.

Comment 6: (page 6) "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."

(page 8) "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."

(page 8) "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."

Response 6: OEHHA evaluated in vitro assays, including ToxCast assays to see whether any information could be gleaned relative to interactions with receptors that might have a relationship with the nervous system. This is described in Chapter 4 and Appendix A. We agree that it is difficult to make conclusive statements about any impacts on children's behavior from these in vitro studies. However, some of the results of the in vitro and HTS assays, as the reviewer notes, are relevant to potential mechanisms of action of the synthetic food dyes on children's behavior.

We note that, while transient effects on neuronal receptors may be reversible, the impacts on a child wo has repeated exposures to chemicals causing these transient effects is significant in terms of that child's behavior and resulting impacts.

Comment 7: "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 veins of conscious rats, when radioactivity was detected in 14 brain regions."

Response 7: We agree that there are areas of the brain where the blood-brain barrier is not contiguous and thus circulating chemicals may enter the brain in these locations. Thus, passive transfer across regions that lack the blood-brain barrier is one way that absorbed dyes may reach the brain.

Comment 8: (page 9) "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 review. Sensitive areas of brain function included regulation of activity, anxiety and exploration in a novel environment, and persistence in the forced swim test."

"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."

Response 8: We agree that these studies found effects on the activity of the animals, and view these as important animal studies in the hazard identification of the synthetic food dyes as having the potential to impact neurobehavior in humans.

Responses to Comments from Peer Reviewer Emily S. Barrett, Ph.D.
Associate Professor of Biostatistics and Epidemiology at the Rutgers School of Public Health and the Environmental and Occupational Health Sciences Institute, Piscataway, NJ

Comment 1: "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 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."

Response 1: Thank you for the thorough review of the epidemiology section of the report. All the elements of a PECO statement are provided in our inclusion criteria: the population (P) is children ≤ 19 years old (inclusion criteria #5, page 30); the exposure (E) is synthetic food dyes (inclusion criteria #3); the comparator are children exposed to placebo (inclusion criteria #6); and the outcomes are neurobehavioral effects related to hyperactivity or inattention (inclusion criteria #4).

As regards a more comprehensive list of outcomes, the search terms used, which are available in Chapter 1, section 1.3.3, were meant to be encompassing. The search terms covered neurotoxicity, neurodevelopmental disorders, neurocognitive, neurologic, neuropsychiatric, and so on, as well as conditions such as ADHD, autism spectrum disorder and general behavioral disorders. The studies we found using the literature search terms and ultimately reviewed focused on a number of measures of hyperactivity and inattention. In addition, a few papers reported on irritability, restlessness, and sleeplessness.

Comment 2: "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."

Response 2: The quality criteria we used were based on the NTP OHAT Risk of Bias tool, and this has now been mentioned in the report. The availability of the quality assessment tools mentioned by the reviewer does not preclude researchers and reviewers from developing and using other better, more specific tools for assessing study quality. The major advantage of the set of criteria we used is that they are specific for randomized clinical trials (RCT) of artificial food dyes and childhood neurobehavior. We developed our criteria by first starting with the National Toxicology Program's OHAT Risk of Bias Rating Tool (NTP Office of Health Assessment and Translation, 2019). We then modified (i.e., improved) this tool by making it more specific to the studies being reviewed. These modifications were based on basic epidemiologic principles (e.g., Rothman and Greenland, Modern Epidemiology, 1998), principles of good RTC design (Hulley and Cummings, Designing Clinical Research, 2013), a preliminary review of the relevant studies, issues raised in a public symposium sponsored by our agency, and concerns raised in critical review articles or reports. The reviewer did not identify any quality criteria that we should have used but did not, or any quality criteria that we used but shouldn't

have. We have now added a more thorough description of the process we used to develop the set of quality criteria used in our report.

Regarding sample size, the reviewer is correct: although sample size was not the sole basis for our inclusions, it was one of many elements we considered in making our conclusions. We are not aware of any broad consensus or any scientific rationale for removing studies with few participants. Studies with small sample sizes have had important impacts on public health in the past (e.g., thalidomide and birth defects, vitamin C and scurvy, diacetyl and bronchiolitis obliterans).

Comment 3: "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?"

Response 3: The cut off for a high dose was approximately the median dose in the clinical dosing studies. This is now mentioned in the report (page 31).

A washout period >2 days was considered adequate. This is now mentioned in the report (page 31).

Comment 4: "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."

Response 4: The reviewer is correct that we could not describe which food dyes were responsible for the effects observed in children because the majority of studies used a mixture. We note in the report that most studies used a combination of agents (page 46). Yellow No. 5 is the one dye that has been studied individually in humans. We have discussion of those studies in our document. We also reviewed animal studies of individual dyes, provide discussion of

those results and use the animal toxicology study results as part of the hazard identification. As well, children are exposed to more than one dye at a time. Thus, mixture studies reflect real world exposures.

Comment 5: "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."

Response 5: As we discuss above and in the report, there is fairly extensive heterogeneity in study designs, study populations, dosing regimens, outcome assessments, and many other factors, across the studies we identified. This heterogeneity limits the value of direct comparisons across studies. Lok et al. (2013) appears to have been designed, at least partially, to replicate McCann et al. (2007), although on the surface the results of these two studies seemed to differ. Because of this, we felt it was important to provide a detailed description of the possible reasons why these results differed. We are not aware of any other two studies where one was meant to replicate the other, but the results differed dramatically. Given all of this, we feel the approach we took in this report, which examines the literature as a whole, provides a much more comprehensive and valuable review of the literature.

Comment 6: "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."

Response 6: As regards the finding that polymorphisms in a histamine degradation gene, one study (Stevenson et al., 2010) found that polymorphisms for histamine breakdown (histamine N-methyl transferase gene) are associated with increased response of children to dyes in the dye mixture studies. There is growing recognition that histamine is an important neurotransmitter in the brain. As reviewed in Scammel et al. (2019), histaminergic neurons are located in the tuberomaxillary nucleus and innervate other regions of the brain including the cortex, thalamus, and other wake-promoting areas. Histaminergic neurons modulate the function of neurons that make dopamine, serotonin, glutamate, GABA, norepinephrine, and acetylcholine. Thus, histamine can have a wide range of effects on brain function. Histamine and serotonin systems are intertwined biologically (Best et al. 2017), providing a link between the Stevenson study in humans and serotonin mechanism studies in animals that we reviewed.

Histamine N-methyl transferase methylates histamine in the extracellular space following release resulting in inactivation of histamine. A number of polymorphisms in the gene have been identified leading to either greater or lesser activity of the enzyme, and several have been associated with specific neurological diseases (Yoshikawa et al, 2019). Thus, the finding by Stevenson et al. (2010) that HNMT polymorphisms modulate the effect of food dyes on activity adds to the evidence that the food dyes can impact brain function, and in this case behavior.

We have now added some of this information to the discussion of histamine in Chapter 4, and added the need to determine genetic determinants of susceptibility to the future research needs section in Chapter 8.

Comment 7: "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."

Response 7: Comment noted. We appreciate the recognition that this review was broader than previous reviews by other authoritative bodies.

Comment 8: "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)."

Response 8: We agree that more study would help clarify the effects of the food dyes on children, for example, by age and gender. However, we think there is enough evidence at this point to consider that food dyes adversely impact neurobehavior in some children. We have added more to our section on research needs in chapter 8 covering the above points.

Comment 9: "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."

Response 9: We agree. Given the widespread interest on the potential health effects of synthetic food additives, it seems somewhat unlikely that a number of large well-conducted clinical trials would remain unpublished.

Comments on 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."*

Comment 10: "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."

Response 10: Regarding the estimate of food dye intake from OTC syrups for fever, allergy/colds, there was an error in the publication we utilized for measurement of dyes in this class of OTC medications that got incorporated into our estimates of exposure from the OTC

syrups. This error was initially raised by scientists working with the Consumer Healthcare Products Association. We confirmed the inaccuracies in our initial color additive exposure estimates for children from cough/cold/allergy and pain relieve/fever reducer syrups, and have made changes to the text and tables of the report to correct the exposure estimates. With one exception, the corrected values are all 100-times lower than initially reported and none of the estimated exposures from the OTC syrups exceed the US FDA or JEFCA ADIs. The error did not affect exposure estimates from tablets or gummies.

Comment 11: "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."

Response 11: The analysis of food dye content in selected pharmaceuticals consumed by pregnant women and children has now been published in the open access peer reviewed literature by the Mitchell lab who did the work for OEHHA. The citation is: Certified Food Dyes in Over the Counter Medicines and Supplements Marketed for Children and Pregnant Women. Lehmkuhler AL, Miller MD, Bradman A, Castorina R, Mitchell AE. Food Chem Toxicol. 2020 Sep;143:111499. As far as we could determine, this is the first publication of quantitative data on food dye content in pharmaceuticals in the open literature.

Comment 12: "In the new analysis, the researchers linked 2015-2016 NHANES demographic and 2day 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."

Response 12: The reviewer is correct to point out that FD&C Green No. 3 was consumed least frequently of all the dyes considered in this assessment. In response to her comments, we have added this point to the report in three places:

1) To our discussion of FD&C dye intake estimates results (mg/person/day) on Page 191:

*"… **FD&C Green No. 3 was by far the least frequently consumed dye with the** lowest median and 95th percentile single and average 2-day intake estimates (See Appendix F Table A3)."*

2) To our discussion of Figures 6.3a-b on Page 196:

6.4.1.3 FD&C Green No. 3

"We calculated single- and two-day FD&C Green No. 3 exposure estimates (mg/kg/day) based on the typical- and high-exposure scenarios for pregnant women, women of childbearing years, and children age 0 to 18 years.

*The median and 95th percentile typical- and high-exposure scenario estimates for single-day exposures (mg/kg/day) are shown in Figures 6.3a and 6.3b, respectively. Children 0 to 18 years old tended to have higher single-day FD&C Green No. 3 exposures compared to women. **FD&C Green No. 3 was infrequently consumed by women and children, and as a result, the typical- and high-exposure scenario median and 95th percentile estimates were the same**."*

3) To our 6.4.1.8 Food dye exposure summary on Page 206:

6.4.1.8 Food dye exposure summary

*"Tables 6.6 – 6.12 provide the exposure estimates for each of the seven FD&C dyes that we assessed. Overall, children's FD&C food color exposure estimates (mg/kg/day) tended to be higher compared to adult women. Among the food dyes, the highest exposures were to Red No. 40 followed by Yellow No. 5 and Yellow No. 6 (Tables 6.6-6.12). **The lowest median and 95th percentile single and average 2-day exposure estimates were for Green No. 3**."*

Comment 13: "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."

Response 13: We agree with the reviewer's comment. In response, we have revised our discussion on Page 223 of the socioeconomic differences in total food dye consumption (mg/kg/day) to make clear that the statistical analyses (Pearson correlations and T-tests) we performed to investigate univariate associations between total food dye exposure estimates and women's and children's poverty levels, race/ethnicity and women's level of education are preliminary results rather than definitive relationships. The changes to Page 223 are noted below:

6.7.4 Summary

"Our analysis of socioeconomic determinants of food dye exposure demonstrate higher intake among non-Hispanic Black participants compared to non-Hispanic whites. The data suggest higher exposure in lower income families with less education, although the differences are not statistically significant. ***These results are based on univariate analyses and should be considered preliminary and hypothesis generating, rather than definitive.***

Another *limitation of this analysis is that the availability of food products in different neighborhoods in the US may vary, impacting exposure. For example, markets in "food deserts" or lower income communities might not carry the same range of products available in more affluent communities, limiting choices. Also, some US supermarket chains, often more expensive, have explicit policies prohibiting sales of foods with synthetic food coloring. Thus, consumers without access to these stores may have higher exposure because their neighborhood markets are more likely to sell foods containing synthetic food dyes even if they are purchasing the same general food categories as consumers in other neighborhoods. Thus, differences in exposure associated with socioeconomic variables may, in part, be due to food systems that unevenly distribute synthetic food dye-containing products into some communities."*

Comment 14: "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."

Response 14: As noted above, there was an error in calculating the amount of food dye in the over-the-counter syrups in the publication used by OEHHA to make these calculations. Those exposure estimates are now corrected and are much lower than before. However, the rest of the exposure estimates are correct.

Response to Public Comments

Response to Comments from the International Association of Color Manufacturers

The International Association of Color Manufacturers (IACM) submitted a comment letter accompanied by two appendices. The first appendix (Appendix A) is a report from Ramboll corporation conducted for IACM that discusses the animal toxicology studies. The other (Appendix B) is a report from ToxStrategies that reviews mechanism related information from ToxCast as well as the peer-reviewed animal toxicology literature, and discusses our evaluation of the high throughput assay systems data and other in vitro data.

The IACM comments are repeated in the initial IACM summary in bullet form and as text, in the Appendix A and in Appendix B. We grouped the comments by area and theme in order to respond to the key points in the comments. We first respond to the bullets and more general comments in the text of the IACM letter, noting the page number from whence we excerpted the quoted comment (section 1 below). This section includes our responses to the comments on the human studies, which were in the body of the IACM letter. We then respond to the specific comments on the animal toxicology studies made by Ramboll in Appendix A (section 2 below), followed by those made by ToxStrategies in Appendix B (section 3 below). The more specific comments repeated in the text of the IACM letter are responded to in these later sections. Finally, we respond to the IACM/ToxStrategies comments on the high throughput assays in Section 4.

1 Responses to IACM summary bullets and body of letter

Comment 1.1 (page1): Several international expert bodies have drawn a different conclusion regarding the potential causal link suggested by OEHHA. Beginning with the European Food Safety Authority (EFSA) in 2008, numerous risk assessment authorities, including the U.S. Food and Drug Administration (FDA) in 2011, have specifically evaluated the clinical evidence purportedly linking consumption of color additives to neurobehavioral effects in children with ADHD and the general population. All of these assessments have concluded that there is no causal relationship. The difference between OEHHA's outcome and those of expert bodies raises serious questions about the validity of OEHHA's investigation, assessment, conclusions, and report.

(and pages 3-5) Regulatory bodies worldwide have evaluated the same studies included in OEHHA's report and all previous assessments have come to a different conclusion than OEHHA. The US FDA, EFSA, and the Joint FAO-WHO Expert Committee Report on Food Additives (JECFA) have all recently conducted risk assessments of and/or reviewed the existing science supporting the continued safe use of food colors. While neurobehavioral endpoints were included in their respective assessments, the pivotal conclusions or recommendations were based on endpoints other than neurobehavioral effects. Even though these risk assessment authorities have reviewed all the available studies on neurobehavioral effects, none have concluded that food colors cause neurobehavioral effects or that the relationship can be established with the available evidence, and therefore do not derive their safety conclusions based on neurobehavioral endpoints.

For example, neurobehavioral effects observed in animal studies have been considered by regulatory and scientific expert bodies. Unlike OEHHA, these expert groups have concluded that these studies or the magnitude of effects within them do not provide evidence that warrant the revision of their respective ADIs. Both JECFA and EFSA concluded that while some animal studies were well-designed and methods were generally well described (Tanaka 2006;

Tanaka et al. 2008), the results did not demonstrate any adverse effects on neurobehavioral development, could not be used in risk assessments, and therefore revision of the existing ADIs was not warranted. Neurobehavioral effects as an endpoint for hazard identification have consistently been determined to be insufficient in a risk assessment or to warrant consideration in the derivation of the ADIs for colors. All the relevant expert risk assessors separately concluded that available data on neurobehavioral effects provided insufficient evidence to base or revise an ADI.

Between 2009 and 2014, EFSA conducted detailed evaluations of six of the nine food colors of interest to OEHHA. None of the six EFSA risk assessments identified neurobehavioral effects as a critical endpoint nor relied on neurobehavioral effects to revise the ADI. EFSA continues to conclude there is insufficient evidence for a causal link between any food color and neurobehavioral outcomes in children. In no case was a neurobehavioral effect considered to sufficiently demonstrate a hazard to warrant establishing an ADI on that basis.

Like EFSA, JECFA based its risk assessments on endpoints other than neurobehavioral effects. As a result of their evaluations, JECFA has maintained or increased ADIs for these food colors and did not consider evidence presented for children's neurobehavioral effects to be sufficient for revising the ADI. Additionally, for all colors, JECFA concluded the McCann et al. (2007) was of limited value, and animal neurobehavioral studies were not considered sufficient or robust enough to be included in the risk assessment. JECFA did identify a one-generation reproductive toxicity study for FD&C Blue No. 1 that evaluated neurobehavioral development in mice (Tanaka et al. 2012) but concluded that the "findings were not robust enough" for purposes of risk assessment.

Lastly, the US FDA Food Advisory Committee (FAC) was asked to evaluate the totality of the evidence and concluded that children's undesired behaviors appeared to result from food intolerance generally, including an individual intolerance to such substances and not a result of neurotoxic properties of the color additives included in the review. During the 2011 FAC meeting, representatives from the FDA stated that none of the agency's ADIs are based on neurological endpoints. However, it was also stated in the FDA's presentations that neurological effects of colors appear to be secondary symptoms of hyperactive behaviors (food intolerance being primary) and there is no indication that colors are directly impacting neurological processes. It was noted that the ADIs of five of the six most commonly used colors were based on two-year animal studies, which included *in utero* exposure as well as dose levels up to the maximum tolerated dose in order to capture the developmental period from the point of conception through end of life at very high dose levels. During examination of those trials, several neurotoxicity screening issues are observed or examined to determine if there is a need to do further specific neurotoxicity testing. The studies did not include specific neurobehavioral testing, but did include clinical observations of lacrimation, clinical observations of behaviors in the normal cage setting.

Response 1.1: In 2011, US FDA asked its Food Advisory Committee (FAC) to consider available relevant data on the possible association between consumption of synthetic color additives in food and hyperactivity in children, and to advise FDA as to what action, if any, is warranted to ensure consumer safety. US FDA concluded that most children have no adverse effects when consuming foods containing color additives, but some evidence suggests that certain children may be sensitive to them.

The FDA states on their web page "*The FDA has reviewed and will continue to examine the effects of color additives on children's behavior. The totality of scientific evidence indicates that most children have no adverse effects when consuming foods containing color additives, but some evidence suggests that certain children may be sensitive to them….*"

As we describe in the Introduction, OEHHA did not limit the review to the question of effects on children diagnosed with ADHD or other behavioral disorders. Rather, OEHHA evaluated the literature to determine whether there might be any effects on behavior of the FD&C batch-certified synthetic food dyes in children in the general population with or without a diagnosis of ADHD. We did not focus solely on effects related to activity and attention, but evaluated the literature for effects on other neurobehavioral impacts as well. In addition, OEHHA evaluated the animal toxicology literature relevant to neurological endpoints; these studies were not emphasized in the 2011 US FDA review.

There is no documentation that is publicly available that the FDA (or JECFA or EFSA for that matter) reviewed the animal literature, as thoroughly as did OEHHA. As can be seen in Tables 3.9 and 3.10 in the OEHHA report, there were a number of animal studies with statistically significant changes in behavior following administration of individual food dyes as well as mixtures of food dyes. Neurobehavioral effects were observed for Yellow No. 5, Yellow No. 6, Red No. 3, Red No. 40, and Blue No. 1. In the mixture studies, neurobehavioral effects were also observed. OEHHA agrees that more studies would be useful; however, the data cannot be simply dismissed. There were statistically significant neurobehavioral effects, for some dyes in multiple studies. As regards consistency of effect, the studies used different protocols, dosing regimens, time to measuring effects and measured different domains of neurobehavior, thus precluding evaluation of consistency. Several well-conducted studies were published after the reviews of USFDA and in some cases after the updates from JECFA. Additional studies have been published since then, both in vivo and in vitro. Finally, we reviewed newer data relevant to mechanisms of action of potential neurobehavioral or neurotoxic effects of the food dyes. The authoritative bodies cited in the comment did not review and integrate all the mechanism, animal toxicology and human studies on the topic of neurotoxicity.

In short, OEHHA did a thorough review of more scientific information than any of the international bodies noted in the comment. It also should be noted that the EU requires that foods with azo dyes be labeled with a warning “may have an adverse effect on activity and attention in children”.
(https://www.europarl.europa.eu/sides/getDoc.do?language=EN&type=IM-PRESS&reference=20080707IPR33563)

As far as development of the ADIs, the ADIs are largely based on very old studies, published decades ago, and are not based on neurological outcomes. The studies that were the basis of the ADI, as noted in the comment, did not measure neurobehavioral effects. As well, the investigators of these older studies performed superficial screening for any neurotoxic effects. Many of the studies used by FDA to develop the ADIs are not publicly available in any detail to conduct a formal study quality review, as IACM contractors provide in Appendix A using the ToxR Tool for publicly available published studies. Given the time in which they were conducted (1950s and 60s; some in the 1980s), they would likely not do well in such a review. The present ADIs were based on old studies not published in any detail that lacked a thorough evaluation of neurobehavioral measures, did not measure outcomes that one would measure in a modern study assessing neurobehavioral effects, and could not be evaluated for study quality. Many of the studies cited by OEHHA found effects at the level of the NOAELs from these old studies used as a basis for the ADI, and in a few cases at the level of the ADI. IACM’s own contractors judged most (11/14) of the individual dye studies OEHHA cites as “reliable without restrictions” or “reliable with restrictions” using the ToxR Tool and note a number of them scored high in their application of the ARRIVE guidelines (see Appendix A, Table A1).

Most importantly, there is substantial evidence from human studies that the food dyes impact

behavior in some children. As noted above, FDA acknowledges this. While there was discussion that the neurobehavioral effects might be due to some "hypersensitivity", that does not eliminate the concern, and no evidence that the effect is from an allergy type phenomenon is cited for this theory. OEHHA presents a thorough systematic review of the literature in chapter 2 of the OEHHA report, including a comprehensive study quality evaluation. The clinical trial studies in children, where children were first given a diet free of food dyes and then provided food dye challenge, and a meta-analysis published in 2012 (Nigg et al, 2012) support an effect of food dyes on children's behavior. One study (Stevenson et al., 2010) found that polymorphisms for histamine breakdown (histamine N-methyl transferase gene) are associated with children responding to dyes in the dye mixture studies. Interestingly, recent scientific studies have shown that histamine is a neurotransmitter and histaminergic neurons interact with many other structures in the brain including those involved in behavior (Yoshikawa et al, 2019; Scammel et al., 2019). OEHHA's report states (in Section 2.8), after a thorough evaluation of the human clinical trials, "Based on the extent of the positive findings reported, and the fact that we could not convincingly or consistently attribute these positive findings to errors in study design or other bias, we conclude that the current human epidemiologic evidence supports a relationship between food dye exposure and adverse behavioral outcomes in children."

Comment 1.2 (page 2): "Without identifying an apparent neurobehavioral hazard for food colors, it is impossible to conduct a risk assessment for that endpoint for food colors. OEHHA's analysis included reported neurobehavioral impacts in human trials, but significant limitations in methodology reduce these studies' value."

Response 1.2: Contrary to what is implied in the comment, OEHHA did identify a neurobehavioral hazard as presented in the assessment in chapter 5. The hazard identification is based on studies from multiple data streams, including clinical trials conducted in children, in vivo animal studies, in vitro mechanistic studies, with some albeit limited support from high throughput assay systems. Specific responses to comments critical of the human and animal studies are detailed below.

Comment 1.3 (page 2): OEHHA's draft report often emphasizes results from select *in vitro* studies that help support a presumed conclusion (i.e., that color additives affect behavior). At the same time, OEHHA dismisses other *in vitro* and *in vivo* data (e.g., Lok et al., 2013) that indicate a lack of evidence for neurobehavioral impacts.

Response 1.3: OEHHA did not have a "presumed conclusion". Rather, we conducted a thorough review and analysis of all available information. It was after this thorough analysis that OEHHA drew conclusions. The *in vitro* data are discussed in the report primarily in Chapter 4.

As well, we discuss Lok et al (2013) in Chapter 2 of the report citing a number of problems with the study. See also the response to human studies criticisms below.

Comment 1.4 (page 2): There is an overall lack of connection between mechanistic data/experimental animal behavioral studies and human disease outcomes in OEHHA's report.

(page 4) "If OEHHA's assertions that all synthetic colors cause neurobehavioral effects were accurate, which the evidence does not support, then there must be realistic and shared molecular mechanisms that would provide some explanation for these effects to be caused by all such colors. The allowed (certified) color additives in the U.S. include azo dyes, a triarylmethane derivative, an indigotine derivative, and an iodofluorescein derivative. These

chemical classes are significantly distinct in chemical structure. Therefore, ascribing behavioral effects to all FD&C colors via a unified mechanism is not scientifically supported, is not addressed nor demonstrated in OEHHA's draft report, and should be carefully questioned."

Response 1.4: OEHHA developed a comprehensive review of available scientific data around the question of whether synthetic food dyes could impact children's behavior. Our conclusion that available data support that synthetic food dyes can impact behavior in some children is well-supported. We don't know which of the of the food dyes are the most important actors, but there are more data available for Yellow No. 5, Yellow No.6, Red No.3 and Red No. 40, this may be in part because they are better studied. We agree that the mechanism for any of the food dyes is not clear and do not purport that the mechanism is the same for all of the dyes causing the observed effects in children. It is not often possible to make absolutely clear connections between results from animal studies and human studies particularly with respect to behavior. However, a number of the animal study endpoints that were observed, including hyperactive behaviors, can be viewed as in the same domain as increased activity in children following exposure to the food dyes. Given the complex contributors and mechanisms that contribute to neurobehavioral outcomes there is no reason why individual chemicals would require action through identical pathways in order to cause similar impacts.

We address comments on our analysis of the ToxCast data below. OEHHA notes in our conclusions that while we cannot ascertain a clear mechanism for any of the dyes at this point, some of the mechanistic data provides support for effects of some of the dyes on neurobehavior.

Comment 1.5 (page 2): In drawing its conclusions, OEHHA gives significant weight to non-guidance studies where weak statistical analysis is used to accentuate inconsistent signals. OEHHA also draws conclusions from "noise" in animal or *in vitro* studies and prioritizes such findings despite overwhelming evidence that supports a conclusion of no effect. Conversely, a lack of consistent results among studies generally leads to a weight-of evidence conclusion that an identifiable hazard does not exist.

Response 1. 5: OEHHA conducted a thorough review of the literature and a review of all animal studies obtained. We reviewed all of the clinical trials in children that met our inclusion criteria. We, in fact, did a very comprehensive review of the studies, including of the statistical analyses used and we note where the statistical analysis was flawed. We derive our conclusions from human and animal studies with sound statistical analyses, not "noise". The clinical studies in children, with appropriate and positive statistical findings support an effect of synthetic food dyes on behavior. We could not find a strong basis for dismissing these studies. As regards consistency of effect, the animal studies used different protocols, dosing regimens, time to measuring effects and measured different domains of neurobehavior, thus precluding evaluation of consistency. Studies used by FDA and other bodies as the basis of their ADIs are many decades old and do not reflect the state of the art in either "guideline" neurotoxicology studies as we know them today or non-guideline neurotoxicology studies published in scientific journals. Note that because a study is a "guideline" study does not in and of itself mean it is a good study, nor that it measures outcomes appropriate to the question of whether food dyes impact neurobehavior. Further, there is not "overwhelming evidence that supports a conclusion of no effect". The clinical studies in children, with appropriate and positive statistical findings support an effect of synthetic food dyes ion behavior. We could not

find a strong basis for dismissing these studies. The animal studies used as the basis of ADIs did not test for the types of effects observed in human studies, or any other specific neurological endpoints. As regards consistency of effect, the studies used different protocols, dosing regimens, time to measuring effects and measured different domains of neurobehavior, thus precluding evaluation of consistency.

Comment 1.6: (page 2) "The majority of meta-analyses and systematic reviews of those meta-analyses published in the last 5-7 years have concluded that dietary intervention methods, including diet- restriction approaches (including color restricting) and those that are pro-nutrient, do not significantly alter children's behavior. These conclusions do not support an association between food colors and neurobehavioral endpoints and should be appropriately considered within OEHHA's analysis and report."

(page 4) "Additionally, as other regulatory bodies have noted, the human studies purporting an association between food color consumption and adverse behavior all suffer from protocol limitations, demonstrating statistically insignificant or inconsistent associations. It is not clear how OEHHA reviewed the same studies as other expert bodies and drew such different conclusions. Other regulatory bodies have repeatedly concluded that there is no consistent evidence of an association between food colors and neurobehavioral effects, much less a causal relationship."

Response 1.6: Most of the dietary intervention studies are either not specifically focused on artificial food dyes, involve reviews of potential ADHD treatments, and/or do not include full evaluations of causal inference.

Comment 1.7: (page 2) "The OEHHA report has a significant flaw regarding its inclusion of studies. While the report suggests that it has taken a systematic approach in its literature search and review, it does not describe the criteria used to qualify or exclude studies. This leaves the impression that OEHHA's weighting of studies in drawing conclusions is either arbitrary, selective for those that fit a narrative, or both."

"The reliability (i.e., quality of methods and reporting) of the studies reviewed and included in the OEHHA report was evaluated only for human epidemiology studies, while study reliability was not formally reviewed for experimental animal studies or mechanistic study data."

(page 8) "The animal toxicity studies that were considered in the OEHHA draft report were not evaluated by OEHHA using any type of published methods for study quality or reliability. OEHHA reports changes in all neurobehavioral outcomes without consideration of study quality, design, consistency in the test methods employed, or interpretation of findings across studies"…"Types of behavioral tests (behavior ontogeny, motor activity [including habituation], motor and sensory function, or learning and memory) were scored using different tools/procedures (automated versus manual, real-time versus post-test, single versus repeat testing of animals) and varied across studies. All statistically significant outcomes appeared to be incorporated into the OEHHA summary of animal toxicology without evaluation or discussion regarding the lack of consistency of effects across studies, between sexes or life stages within the same study, or of the observed animal-to-animal variation in effects upon exposure to colors versus those in the control groups."

Response 1.7: The comment that we did not look at the available information using a systematic review protocol is incorrect For animal studies, we conducted a thorough review of the literature and essentially retrieved and reviewed all articles, as there is not a large

database for any of the dyes. We used criteria similar to those in the ARRIVE guidelines for evaluating study quality and in our large tables of animal studies, we provide information that coincide with the areas included in the ARRIVE guidelines.

We did not use the ToxR or SciRAP software program to include/exclude studies. In addition, we did not rely on titles and abstracts as an initial screen, which can lead to dismissing studies that should be considered. In contrast, we read the entire papers before deciding whether or not the paper was providing relevant information. In addition, the IACM contractors focused on dose-response assessment and not whether the study provides information on hazard. This overly constrained approach to literature review is not appropriate for conducting a thorough review of the scientific literature to determine whether there is potential for neurobehavioral effects of the synthetic food dyes in children.

As regards consistency of effect, the studies used different protocols, dosing regimens, time to measuring effects and measured different domains of neurobehavior, thus precluding evaluation of consistency. Our response to comment 2.2 below goes into more detail on the issue of consistency. Finally, we discuss the statistics for each animal study and note where there were limitations.

Comment 1.8: (page 2) "Studies of mixtures of food colors are not appropriate for hazard identification. They do not allow the identification of specific food colors that might pose a hazard, if such a hazard exists. Additionally, many of the studies include color additives within the mixtures that are not approved for use in the United States nor within the scope of OEHHA's review. In fact, by considering combinations of colors, OEHHA has, in many cases, asserted effects for color additives that likely have no contribution to the identified hazard, if such a hazard exists at all within the study."

Response 1.8: OEHHA agrees that studies of mixtures present a problem with attribution of effects to any one dye. OEHHA states that, "Most studies involved administering multiple dyes at the same time so no single offending agent could be identified" (page 46). Regardless these studies can still be used to evaluate whether the broad class of artificial food dyes as a whole might contain at least some agents that cause adverse neurobehavioral effects. Moreover, children are not generally exposed to one dye at a time, but rather to multiple dyes simultaneously. As such, mixture studies reflect real world exposures. Indeed, almost all of the dyes were studied in humans as mixtures. The exception is tartrazine. However, the larger well-conducted mixtures studies showed effects from the dye mixtures (see Table 2.1). In addition, there are animal toxicology studies of individual dyes that demonstrate effects on behavior and the results of the animal studies are part of our hazard identification.

Comment 1.9 (page 3): "OEHHA prematurely concludes that food colors may cause or exacerbate neurobehavioral problems in some children and suggests that current acceptable daily intakes (ADIs) are not protective of children and should be lowered."

(page 4) "OEHHA's conclusion on page 20 of the draft report that the current FDA ADIs are not adequately protective of children, is not based on sound science. The ADI represents a conservative upper daily intake that is not expected to result in an adverse effect in the most sensitive individuals regarding general and organ-specific toxicity, including reproductive, developmental, neurotoxicity, genotoxicity, and other forms of toxicity. Authoritative bodies develop ADIs based on points of departure, frequently the no observable adverse effect level (NOAEL) or Benchmark Dose level (BMDL). The ADIs for FD&C colors account for the most sensitive 10 percent of individuals, assuming that humans are 10 times more sensitive than

the most sensitive test species. As noted previously, while the animal studies examining the potential effects of food colors on neurobehavioral endpoints were included in the recent regulatory safety reviews of EFSA and JECFA, none of those studies presented evidence of adverse effects that would warrant revision of the established ADIs."

Response 1.9: Our conclusion that "the scientific literature indicates that synthetic food dyes may cause or exacerbate neurobehavioral problems in some children" is not premature, but rather is based on our thorough evaluation of all available information. Data from multiple evidence streams, including epidemiology, animal neurotoxicology, and some *in vitro* and high throughput assays, support this finding.

Studies used by FDA and other bodies as the basis of their ADIs are many decades old and do not reflect the state of the art for either "guideline" neurotoxicology studies as we know them today or non-guideline neurotoxicology studies published in scientific journals. As noted by the commenters, these studies did not measure neurobehavioral effects. In our review, many animal studies find neurobehavioral or other neurological effects at the NOAELS used by FDA for the ADIs (see Chapter 7, section 7.3 and Tables 7.2a and 7.2b). Given that neurobehavioral effects have been observed in clinical trials in children and that animal studies have shown effects at levels lower than the NOAELS used by the FDA to derive the ADI, it is not possible to conclude that the ADIs are therefore protective of children. In the available reviews, JECFA provided a brief statement concerning some individual studies but did not review them in the context of previous studies and broader neurotoxicology literature, all the children's studies and children's exposure data as was done in the OEHHA review.

Comment 1.10: (page 3) However, OEHHA does not provide any conclusive evidence of causality that would warrants any risk management actions, including lowering ADIs.

Response 1.10: OEHHA's report does not delve into risk management options. The epidemiologic literature we emphasize is comprised of human clinical trials that used a cross-over blinded design. These can be interpreted to provide strong evidence of causality by the nature of their design. We present a thorough evaluation of the human, animal, and in vitro studies available and synthesize the information. Together the available information supports that food dyes have neurobehavioral effects in some children. This conclusion is in fact not that different from the FDA's conclusion that some children appear to be sensitive to food dyes. Also, our conclusion that the totality of the data supports that food dyes have neurobehavioral effects in children is consistent with the EU's warning label, which requires that foods containing azo dyes be labeled with the warning "may have an adverse effect on activity and attention in children". Further, we make obvious that the existing ADIs are based on decades old animal toxicology studies that were not capable of measuring the types of neurobehavioral responses observed in children. Thus, one cannot claim that the ADIs are protective of neurobehavioral effects in children.

Comment 1.11: (**page 5**) " Overall, design problems exist across epidemiologic studies that attempted to assess the causal relationship between food dyes and potential neurobehavioral concerns such as ADHD in children. In fact, OEHHA stated (page 82, Section 2.4.2), "It appears that the recruitment of participants in most of the studies we reviewed involved convenience sampling, and few studies provided enough information for us to calculate or estimate participation rates. Because of this, we could not use overall participation rates or other subject selection criteria as an indicator of study quality."

Response 1.11: OEHHA evaluated the possibility that convenience sampling would cause major bias. Given that randomized double blinded cross-over designs were used in most of the epidemiologic studies reviewed, OEHHA has concluded that major bias caused by convenience sampling was unlikely (page 44).

Comment 1.12: (**page 5**) "Other shortcomings in study design included proper assessment of the exposure (e.g., exposure to different dyes and/or mixtures of multiple dyes, different purity standards across countries); temporality (adequate washout period, timing of exposure vs. testing, whether the exposure to food dye potentially results in either transient or long-term effects);consideration of bias, chance, and confounding issues (e.g. other potential causes of neurobehavioral problems); and/or problems with scoring methods between different assessors (teachers, researchers, or parents using different metrics or tests)."

Response 1.12: Evaluations of exposure assessment, temporality, bias, chance, confounding and outcome assessment are already provided (pages 40-70). Currently, it is unknown whether the effects identified in some of the studies OEHHA reviewed are transient or long-term. Regardless, even if the effects are transient, they could have long-term impacts in children who are exposed multiple times a day over many years (which many children are). Because most of the human studies OEHHA reviewed used cross-over designs, differences in purity standards would most likely bias results to the null, not cause the positive associations seen in many of these studies.

Comment 1.13: (**page 6-7**) "OEHHA noted in its report (page 29, Section 2.1) that it did not perform a full meta-analysis due to the "high quality" publication of Nigg et al. (2012). This meta-analysis identified 24 studies published from 1976 through February 2011 and evaluated behavioral effects (relative to inattention and hyperactivity) and color additives. Of the 24 studies included in Nigg et al. (2012), only 11 studies, including McCann et al. (2007), evaluated hyperactive children. The authors noted a wide variation in responders between studies. They reported that some children in the reviewed studies saw a reduction in ADHD symptoms on restriction diets. In focusing exclusively on Nigg et al. (2012), the OEHHA report notably does not consider many other meta-analyses and systematic reviews, including Sonuga-Barke et al. (2013), Stevenson et al. (2014) or Pelsser et al. (2017). Going further back into the literature, even a study by Kavale and Forness (1983), an earlier meta-analysis of studies that evaluated several elimination diets, including the Feingold diet, and challenge trials with color additives found no statistically significant association between color additives and hyperactivity in children. Further, no clear criteria were established in the OEHHA report to conclude that the Nigg et al. (2012) analysis is sufficiently high quality to disregard the full weight of evidence, and raises the question why OEHHA emphasized a study finding an effect while minimizing those that did not."

"As highlighted during a meeting of the FDA's Science Board in October 2019 on the topic of Color Additives and Behavioral Effects in Children[3], and emphasized by IACM's February 19, 2019 comments, there are two additional recent publications that are critical for OEHHA to consider in its assessment, a meta-analysis of dietary interventions for ADHD by Sonuga-Barke et al. (2013), and a systematic review by Pelsser et al. (2017)."

"Sonuga-Barke et al. (2013) was published on behalf of the EUNETHYDIS11 European ADHD Guidelines Group, used largely the same dataset as Nigg et al. (2012) and reported similar statistically significant but small (yet clinically insignificant) effects on symptoms of ADHD from ingestion of color additives, but drew slightly different conclusions. The authors identified eight papers that evaluated food color additives which met the authors' criteria for inclusion. The

meta-analysis revealed a statistically significant but weak association between food color restriction and improved behavior. Significantly, when the analysis was limited to (1) four papers that utilized a protocol with low or no pharmacological interventions (because allowing subjects to continue with taking medication to treat ADHD may reduce the ability to detect a potential effect due to food colors) and (2) protocols that were likely blinded, the association between color additives and the behavioral impact was further reduced and ceased to be statistically significant. Pelsser et al. (2017) performed a systematic review of two meta-analyses (Schab, et al., [2004] and Nigg, et al., [2012]) that evaluated the evidence associated with elimination diets for food color additives and ADHD and concluded that the current evidence does not support restriction of food color additives for the treatment of ADHD."

"OEHHA also chose not to consider a second publication by the EUNETHYDIS European ADHD Guidelines Group, Stevenson et al. (2014), which focused only on dietary treatments for ADHD. Stevenson et al. (2014) reviewed three meta-analyses related to ADHD and the purported efficacy of the artificial food color elimination diet (i.e., Schab et al., [2004]; Nigg et al., [2012]; Sonuga-Barke et al., [2013]). Stevenson et al. (2014) concluded that the effect size was too small to be of value and that the patient population that would benefit from a color additive elimination diet remains uncertain. Consistent with previous evaluations, the authors ultimately came to the same conclusions that the methodology used in most of the trials on which the meta-analyses are based were weak, limiting their ability to demonstrate an efficacious treatment for ADHD."

"Taken together, reviews of the clinical trial literature associated with ADHD and the consumption of color additives have produced neither consistent nor strong association between color additive intake and adverse neurobehavioral effects. Furthermore, removal of color additives from the diet has not been demonstrated to be an efficacious treatment of ADHD. None of the studies conducted to date have individually or collectively provided evidence to support the conclusion that an association exists. Moreover, the weak evidence that is detected through systematic review and meta-analysis has been inconsistent and likely the product of subjective diagnostic criteria."

Response 1.13: A large portion of the review by Sonuga-Barke et al. (2013) involved topics that were not directly related to artificial food dyes. This includes research on free fatty acid supplementation, neuro-feedback, cognitive training, and other unrelated behavioral interventions. While the Sonuga-Barke et al. (2013) review included seven studies of elimination diets, these were not specific for artificial food dyes. That is, they included studies of antigenic foods, specific provoking foods, general elimination diets, and oligoantigenic diets. Because these seven studies involved the elimination of multiple foods containing many different chemical agents, it is difficult to ascribe any observed effects seen in these studies directly to artificial food dyes.

Sonuga-Barke et al (2013) also included eight clinical trials of artificial food dye exclusion. However, the comment is incorrect regarding the subgroup analysis involving these studies: when the authors restricted their analysis to the four studies that had no or low medication use, the effect size was not reduced. Rather, it remained unchanged: with a standardized mean difference of 0.32 in both the full and in the subgroup analyses. This effect size is fairly large when one considers the large numbers of people who are exposed to artificial food dyes. It should also be noted that the subgroup analysis referred to by the commenter had nothing to do with blinding. The confidence interval was wider in this subgroup analysis compared to the analysis involving all studies, but that is expected given the smaller number of studies. It

should also be noted that while the Sonuga-Barke et al. (2013) review only included eight studies of food dye elimination, OEHHA's final analysis included 27 human studies, the very large majority of which were randomized cross-over, double blinded, and placebo-controlled designs. Many of these were clinical dosing trials, where known doses of artificial dyes or a placebo were given. This is a much more direct and informative design than that used by the elimination studies included in the Sonuga-Barke et al. (2013) review.

With regards to Pessler et al. (2017), the very large majority of this review involved studies of broad dietary interventions (e.g., diets eliminating many different foods), or interventions that did not involve artificial food dyes (e.g., poly-unsaturated fatty acid supplementation). The authors of this review concluded that the effect sizes in the meta-analyses of artificial food dyes were, "...too small to contribute to ADHD treatment". However, this conclusion was based on estimates of ***average*** effect size. It did not account for the likelihood that some individuals will have greater responses than others. More importantly, this conclusion was for treatment efficacy, not for causation. In fact, the authors stated that the effect sizes in the meta-analyses it reviewed were, "…too large to dismiss." As mentioned in OEHHA's report, even small effect sizes can have large impacts when examined on a population basis. This is especially true if the exposure of interest is widespread, which is the case for artificial food dyes. The Pelsser et al. (2017) authors also stated that, "our conclusions are based on effect sizes rather than on p-values", and all of the effect sizes they provided for the Nigg et al. (2012) and Schab et al. (2004) meta-analyses were consistent with an adverse effect of artificial food dyes on childhood behavior. Finally, the Pelsser et al. (2017) review did not provide evaluations of the quality of the underlying studies that went into these meta-analyses, and did not include a thorough discussion of causal inference. This is in contrast to OEHHA's report, which includes an extensive review of the underlying studies, thorough evaluations of bias and confounding, and analyses considering all of the major aspects of causal inference.

As with Pressler et al. (2017), the review by Stevenson et al. (2014) also focuses on ADHD treatment, not causality. Regardless, the authors of this review state, "Artificial food colours exclusion may be beneficial for children thought to be adverse responders to food colour exposure."

Overall, the publications by Sonuga-Barke et al (2013), Pressler et al. (2017), and Stevenson et al. (2014) were reviewed for OEHHA's report. However, because the major focus of these reviews were on broad dietary patterns, unrelated supplements, or on treatment efficacy, and because they did not include detailed evaluations of study quality, bias, confounding, or causation, they were not extensively described in OEHHA's report.

The goal of OEHHA's assessment was to determine the level of evidence supporting whether neurobehavioral effects occur in children as a result of exposure to synthetic food dyes. We focused on this rather than limiting our review to whether or not the clinical diagnosis of ADHD is impacted by exposure to synthetic food dyes or whether removal of exposure was an effective treatment for ADHD. In addition, we included in our considerations the neurobehavioral effects in children in the general population without any identified behavioral problems or diagnosis. The epidemiologic findings from double-blinded dosing studies that children in the general population have identifiable alteration in behavioral measures, along with the findings in the toxicological and mechanistic literature, support our conclusions.

Comment 1.14: (**page 7-8**) "The only challenge study that attempted to replicate the findings

of McCann et al. (2007) in a different population was published by Lok et al. (2013) which, OEHHA notes (page 43, Section 2.7.4), was not included in the Nigg et al. (2012) meta-analysis. Lok et al. (2013) replicated the design of the McCann et al. (2007) study including the use of a randomized double-blind placebo-controlled design with a within-subject cross-over protocol, to assess hyperactivity in eight to nine year-old children in Hong Kong. The authors stated that this study "does not attempt to negate or contest the findings of the Southampton study but to build on this study in a sample of Chinese children because food safety in China is a major public health issue," hence the study adhered to a protocol very similar to that used by McCann et al. (2007). Lok et al. (2013) used the same doses of the same color additives used in McCann et al. (2007) in Mix A (FD&C Yellow No. 6, FD&C Yellow No. 5, carmoisine, and ponceau 4R) with a few differences such as attempting to obtain a study sample including children from a wide spectrum of socioeconomic backgrounds. More notably, Lok et al. (2013) excluded children with ADHD and currently treated with medication, diagnosed with diabetes, with phenylketonuria (in which interference with diet can have adverse health effects), and/or other mental health problems (e.g., learning disability, Down syndrome) from the study population. These exclusion criteria were not applied in the McCann et al. (2007) study. Other differences include administration of treatment given as capsules instead of juice; timing of administration was specified to be in the morning before school rather than anytime of the day; the preservative sodium benzoate was not included in the same treatment as food colors but was tested separately; and assessment was based on teachers and parents only but with no independent assessor. Lastly, Lok et al. (2013) assigned two types of scores: a) strengths and weaknesses of ADHD symptoms (positive and negative scores) and normal behaviors (SWAN) rating scale (based on the DSM-4 diagnostic criteria) that teachers and parents filled in; and b) child behavior checklist (CBCL), which only the teachers filled in, both of which have been validated with local norms in Hong Kong."

"In contrast to McCann et al. (2007), Lok et al. (2013) did not detect an association between color additive intake and behavior, even though evaluation was also based on parent and teacher assessments. Because Lok et al. (2013) used essentially the same protocol as McCann et al. (2007), IACM recommended in our February 19, 2019 comments that OEHHA carefully review the study protocol and findings from Lok et al. (2013) to assess the lack of reproducibility of McCann et al. (2007)."

"While the results of McCann et al. (2007) were the impetus for a renewed interest in the potential for food color additives to negatively affect behavior in children, regulatory agencies worldwide have dismissed and discounted the findings from McCann et al. (2007) in their own subsequent reviews. Lok et al. (2013) studied a different population using a very similar protocol and larger doses of the food color additives but was unable to reproduce the results, which should raise significant concerns for OEHHA as it has with other regulatory agencies, diminishing the confidence in results from McCann et al. (2007), and be viewed as offering limited value due to the lack of reproducibility."

"Unfortunately, the OEHHA report provides only a cursory review of Lok et al. (2013) and ultimately dismisses the study, seemingly because independent observers and computerized tests like the Conners tests were not used as in McCann et al. (2007), even though the study uses two behavioral metrics based on parent and/or teacher reporting that have been validated in China. It is unclear why OEHHA allows for one psychological test and not the other, when both have been validated as viable tools for clinical assessment. OEHHA also notes that Lok et al. (2013) did not report an association for any outcome metric and proceeds to dismiss the study's validity due to no observed effect."

"Even without considering Lok et al. (2013), reviews of the clinical trial literature associated with ADHD and the consumption of color additives have not produced consistent associations between color additive intake and undesired neurobehavioral symptoms. While there are numerous clinical studies that attempt to investigate the relationship between food colors and potential neurobehavioral effects, none of the studies conducted have succeeded in providing the evidence that would support the conclusion reached by some meta-analyses and OEHHA that an association exists. Unfortunately, OEHHA's report lacks transparency in the criteria used to assess the quality of these studies."

Response 1.14: The commenter highlights a number of major differences between the Lok et al. (2013) and McCann et al. (2007) studies. Because of these differences, Lok et al. (2013) cannot be considered a "replication" of McCann et al. (2007). OEHHA has provided much more than a "cursory" review of the Lok et al. (2013) study. For example, a description of the major aspects of the Lok et al. (2013) study design and a summary of its results are presented in Table 2.1. Multiple study design features and study quality factors for Lok et al. (2013) are listed in Table 2.3a. An extensive discussion of the differences between Lok et al. (2013) and McCann et al. (2007) and several possible reasons why the results of these two studies may differ are presented in Section 2.7.4, where OEHHA lays out the review criteria quite clearly.

Comment 1.15: **(page 7)** "It appears OEHHA has not seriously considered if test articles used in reported studies are comparable to purity as well as impurity standards set forth by regulations for batch-certified food colors that industry uses. Declaring only the purity information with test articles bought from Sigma is not sufficient to assess compliance to batch certified standards. The specific impurity information becomes vital since the minimum purity requirements for food colors under FDA regulations may be lower than the 99.5% material provided by Sigma, and also come with restrictions on specific impurities. Further, widespread accessibility to FDA testing procedures used in batch certification process is limited, and unless the study specifically indicates that the food colors used meet the identity and specifications as regulated by FDA, it is not possible to determine if a test article is comparable to an FDA batch-certified color. This is important contextually when reviewing experimental data because color test articles used in studies are generally not reported to be batch certified under the Federal Food Drug & Cosmetic Act, and many studies lack sufficient information about the purity and extent of critical impurities and reaction by-products in test articles for which maximum limits are set by the regulation of the color additive."

"For example, FD&C Red No 3 under 21 CFR 74.303[2] not only needs to meet the purity requirement of 87%, but contains specifications that restrict the levels of other iodinated compounds like sodium iodide, (≤ 0.4%), triiodoresorcinol (≤ 0.2 %); 2(2',4'-Dihydroxy-3', 5'-diiodobenzoyl) benzoic acid (≤ 0.2 %); monoiodofluoresceins (≤ 1.0 %); and other lower iodinated fluoresceins (≤ 9.0 %). It is vital that the test articles used in studies meet these specification requirements as excess iodide from highly bioavailable impurities like sodium iodide or other iodinated compounds can lead to confounding outcomes. In this case, even with presence of 13% sodium iodide, a test article would still meet FDA purity criteria but not impurity restrictions. This can result in a biological outcome due to this impurity rather than due to the dye. Like FD&C Red No 3, restricted impurities and limits for all FD&C colors can be found in 21 CFR Title 74."

Response 1.15: Our in-depth search of the literature was not confined to studies using the FDA version of the dye. Similar authoritative reviews cited by IACM (US FDA, EFSA, and the

Joint FAO-WHO Expert Committee Report on Food Additives (JECFA)) also did not confine themselves to studies using FDA certified versions of the dyes.

OEHHA did take a serious, in-depth look at the sources of the dyes.

Many of the early animal toxicology studies we reviewed (Sobotka 1977, Vorhees 1983a,b, Shaywitz 1979, Goldenring 1980, Reisen and Rothblat 1986) were conducted in the US and purchased FD&C dyes from Kohnstamme, the same supplier as for studies conducted at FDA during that era (Collins et al. 1976, 1993a, 1993b) and for industry sponsored toxicology studies submitted to FDA in connection with dye certification (Borzelleca 1987a,b).

In general research conducted internationally obtained dyes approved for use in food by their local regulatory agencies.

The six Tanaka et al. studies were conducted at a government sponsored research lab in Japan and used dyes from Tokyo Kasei company. The Tokyo Kasei company supplies certified "lots" of dyes approved for use in foods in Japan (https://www.tcichemicals.com/US/en/c/10111). The lot numbers are provided in the Tanaka publications. Laboratories in Turkey (Doguc et al., 2013,2015) used dyes with Turkish regulations as posted with FAO. In China (Gao et al. 2011) food dyes are also regulated.

Commercial dyes with standardized production and quality assurance, like those sold by Sigma, make it possible for different labs around the world to use the same version of the dye, and for interested parties to independently identify contaminants. The most recent studies (Dalal and Poddar 2009, 2010, Rafati et al. 2017, Noorafsham et al. 2018, Erickson et al. 2014) obtained research quality dyes from Sigma. Sigma also supplies a certificate of analysis for each lot of dye that is sold.

The importance of "critical impurities" is mentioned by IACM in connection with iodinated impurities in FD&C Red 3 which may be more bioavailable than the dye itself. It is important to note that FDA's purity standards for Red No. 3 were established in 1969 along with an ADI based on a single study conducted at FDA. None of the toxicology data for Red No. 3 developed after that time could have been used to establish contaminant limits, which appear to have been derived from GMP (Good Manufacturing Practices). IACM comments did not describe the "biological outcomes" that might be caused by "iodinated compounds" but OEHHA is aware of a literature concerning FD&C Red 3 mediated thyroid cancer. International regulatory agencies have banned or restricted use of Red No. 3. Whether the current contaminant limits provide protection from neurobehavioral toxicology is unknown.

In researching this topic, OEHHA became aware of two studies sponsored by color manufacturers and submitted to FDA (Borzelleca 1987a,b) in which the reported sodium iodide content of FD&C Red No. 3 exceeded the 1969 standard. This raises concern that FDA batch certification process is not always successful in achieving contaminant limits.

Comment 1.16: (page 7) "Furthermore, we note that bioavailability of the colors was generally not addressed in the OEHHA report, except for including absorption and bioavailability as important research needs for future studies. It should be noted, however, that bioavailability is low for all seven colors (EFSA, 2009a,b,c, 2010, 2011, 2014; WHO/FAO, 2017a,b, 2019)."

Response 1.16: Bioavailability and absorption are aspects of ADME reviewed in Section 4.1 of the OEHHA document. All the published articles from the regulatory documents cited in the IACM comment were reviewed in Section 4.1. It was important to review the information on absorption in the context of the different methods used at the time the research was

conducted, and the information on distribution, metabolism and excretion. For example, as discussed in the regulatory reviews, because of the extensive metabolism of these azo dyes in the GI tract, the absorption of bioactive metabolites is more relevant to toxicity than the absorption of parent compound. We also searched and reviewed recent literature because all the literature cited in the regulatory documents was conducted between 1953 and 1986.

OEHHA reviewed the full text of all the original research reports that were publicly available on this topic and, additionally, reviewed JECFA, EFSA and FDA documents for research that may have been reported to these regulatory groups but was not publicly available. Four unpublished articles were reviewed in the regulatory documents (Obrist 1986, White 1970, Guyton and Reno 1975, Guyton and Stanovick 1975); If any quantitative data were presented in regulatory reviews from unpublished studies, we included it. These unpublished studies were submitted directly to the regulatory agencies, were not publicly available, and were not provided with the IACM comments.

Although it may appear that bioavailability of the food dyes is low, this is based on old, very limited data not presented in any detail in published studies. The methods used are old and thus the data on the bioavailability of the food dyes are inadequate. As noted above and in the OEHHA document, Chapter 4, there are some data available, for example on the bioavailability of the azo dye metabolites. Studies have found that Yellow No. 5 undergoes azoreduction in the gut, and the metabolite sulfanilic acid is absorbed and excreted in urine in animals and humans. Sulfanilic acid may be the ultimate toxicant rather than the parent compound. Another example is that, once absorbed, Red No. 3 may reach the brain (detailed in Chapter 4). Following injection of labelled Red No. 3 in the femoral vein of conscious rats (Levitan et al. 1985), radioactivity was detected in 14 brain regions. The authors comment that "significant uptake" could occur with an immature or damaged blood-brain barrier, and in some lower brain regions, including the hypothalamus, not included in the blood-brain barrier. As well, there is evidence of the excretion of de-iodinated erythrosine metabolites in animals. This in itself indicates some absorption of Red No. 3. We call for more research in our concluding chapter on absorption, metabolism, distribution, and excretion of the food dyes and any metabolites in animals and humans. Nonetheless, the findings in the positive animal studies OEHHA reviewed are based on statistical analysis of the outcomes compared to concurrent controls indicating an effect of the dyes. Further, the epidemiological evidence indicates an effect of the dyes whether or not the absorption is low.

Comment 1.17: (page 7) "Statistically significant behavioral results were provided as evidence when pairwise or trend tests were positive, yet without associated specific reference to the importance of litter randomization prior to testing and litter-based statistical analysis."

Response 1.17: Litter-based design or statistical analysis is important when rodent subjects are exposed during pregnancy and lactation by dosing of the dam. OEHHA reviewed these studies and included lack of litter-based design/analysis as a "weakness" in "strengths and weaknesses". Developmental test battery effects from the Tanaka and Sobotka studies were not used by OEHHA to identify study NOAELs and LOAELs because of lack of litter-based design/analysis. We did not find any review for litter-based design/analysis in the Ramboll-generated materials.

Comment 1.18 (page 12) "OEHHA concluded that the Tanaka (2001) study had a no observed adverse effect level (NOAEL) of 28 mg/kg/day which is lower than the FDA NOAEL

of 250 mg/kg/day based on the observation of "distended cecum" in rats, suggesting that the ADI should be protective against neurobehavioral effects."

Response 1.18: If the neurobehavioral NOAEL is 28 mg/kg/d and the FDA NOAEL used to develop the ADI is 250 mg/kg/d, the ADI **would not be protective against neurobehavioral effects**. These effects could occur at doses higher than 28 mg/kg/d but lower than 250 mg/kg/d. (This assumes that the same 100 X uncertainty factor would be used to derive the ADI from either NOAEL).

Comment 1.19: (page 13) "OEHHA concluded that the lowest NOAELs identified of 175 mg/kg/day in mice and 125 mg/kg/day in rats (Gao et al. 2011) were lower than the NOAELs identified by FDA of 500 mg/kg/day in dogs and 1000 mg/kg/day in rats, and 1.5 percent in diet estimated at 2250 mg/kg/day in mice. It is worth noting that the endpoints used by OEHHA to identify NOAELs went beyond neuro-related endpoints, bringing into question the relevancy of the conclusions reached."

"OEHHA stated that a NOAEL without a LOAEL in the same study is not suitable for risk assessment. It is worth noting that a study does not need a LOAEL to have a NOAEL."

Response 1.19: OEHHA was not able to identify the basis for this IACM comment. Behavioral endpoints were used for the identification of the NOAELs in the Gao et al. (2011) study in the OEHHA review:

"In the rat activity test (Open Field), greater activity in terms of both horizontal movement (squares crossed) and vertical activity (rearing) was seen in the mid- and high-dose groups (250, 500 mg/kg/day) compared to controls. The mid- and high-dose (350 and 700 mg/kg/day) interfered with the learning of mice in both the Morris Water Maze and the Step Through Avoidance task. The NOAEL was found at the low dose in both studies (125 mg/kg/day rats, 175 mg/kg/day mice)."

OEHHA agrees that a study does not need to have a LOAEL to have a NOAEL. The OEHHA statement is about suitability for risk assessment.

2. Response to comments on animal toxicology studies

The International Association of Color Manufacturers (IACM) sent a comment letter accompanied by two appendices, one of which (Appendix A) is a report from Ramboll corporation conducted for IACM that discusses the animal toxicology studies. The IACM comments on animal toxicology studies reviewed by OEHHA in IACM's Appendix A (text and Table 1) are reviewed in this section.

2.1 Response to comments derived from Ramboll materials

Introduction to OEHHA responses: With regard to comments based on the Ramboll materials (Appendix A), it is difficult to identify the basis of the comments and conclusions derived from Ramboll materials. A major issue is the citation of a reference, Gentry et al. (submitted). Gentry et al. (submitted) was cited 33 times as support for comments and conclusions in the IACM document. In concluding its introduction to the appendix Ramboll states: "Based on Gentry et al. (submitted), none of the studies evaluated provided strong or consistent evidence of effects." (Appendix A, page 6)

OEHHA is not able to judge the quality or value of Gentry et al.(submitted) as a basis for IACM conclusions. Gentry et al.(submitted) is not publicly available, and no pre-publication version was provided with the comments. Another a report cited in the document (Ramboll 2020) is also not available to the public and was not provided with the comments. Any conclusions based on Gentry et al. or Ramboll (2020) must be considered unsubstantiated.

A second difficulty is the assumption in the Ramboll materials is that the OEHHA review is a quantitative risk assessment of neurobehavioral dye effects in children. As stated in the introduction, the OEHHA assessment:

- Evaluated the literature on human studies relevant to whether behavior is affected in children when they consumed synthetic food dyes.
- Evaluated the literature relevant to neurobehavioral effects in laboratory animals following synthetic food dye exposure.
- Examined information relevant to how synthetic food dyes might exert neurobehavioral effects, including data obtained through high-throughput screening assays (laboratory tests that evaluate the effects of chemicals on cells or biological molecules) conducted by the US Environmental Protection Agency and its federal agency partners.
- Along with collaborators at the University of California's Berkeley and Davis campuses, estimated exposures to each FD&C batch-certified synthetic food dye in general use in the US for children of varying age groups as well as for pregnant women and women of childbearing age.
- Conducted a risk characterization where we present a number of comparisons to gauge whether exposure to food dyes may present risk of neurobehavioral impacts.

The OEHHA review did not conduct dose-response modelling or uncertainty factor selection or culminate in quantitative risk assessment. The term "quantitative risk assessment" did not appear in the OEHHA review but is found 65 times in the IACM comments. For example, the Ramboll Appendix Table 1. is titled "Comparison of Ramboll and OEHHA reviews of Animal Studies Considered in OEHHA (2020) for Quantitative Risk Assessment". Ramboll comments on exclusion of studies because they are not suitable for quantitative risk assessment are irrelevant to the OEHHA review since we did not conduct a quantitative risk assessment (although we comment on whether the FDA ADI could change if the FDA used newer studies to evaluate the ADI).

A third major difficulty is the Ramboll evaluation of the studies on the basis of guidelines for conduct and evaluation of regulatory studies. For example, Ramboll categorizes use of mice in studies as a limitation because rats are the preferred animal model in OECD guidelines for regulatory studies.

OEHHA undertook an exhaustive and in-depth review of all available research. No neurotoxicology or developmental neurotoxicology studies of food dyes have been performed according to FDA or OECD guidelines and submitted to regulatory agencies. Further, such guideline studies would not detect the effects reported in the children's studies reviewed by OEHHA, including immediate effects on behavior after exposure to food dye challenges. A broad range of neurotoxicology and neurobehavioral toxicology studies needs to be reviewed based on current scientific knowledge in these areas, Guideline studies, if available, would still require this type of review.

Despite these difficulties, OEHHA has prepared general responses to IACM comments.

2.1.1. Response to Ramboll comments in Appendix A Table 1 column headed "Ramboll counterpoints to OEHHA arguments" (starts page 7, Appendix A)

The OEHHA arguments referred to in the column header are not identified in this column. However, each study entry in the column begins with a statement that the study did not reach Ramboll criteria for "suitability", "inclusion" or "consideration" for "quantitative assessment of neurobehavioral effects in children". As stated previously, the OEHHA document presents an in-depth scientific review of literature and not a quantitative risk assessment. We do not agree or disagree with the Ramboll conclusion as to whether a particular study should be selected for quantitative risk assessment, but we can respond more generally to the comments supporting that conclusion in the context of the mandate for the OEHHA scientific review.

Comment 2.1 "…the study was limited by low study quality due to the small number of animals per dose group which decreases the statistical power of the results and limits the value of the study results for dose- response."

Response 2.1: This comment, identically worded, is provided for all six of the Tanaka et al. studies (Tanaka 1994,1996,2001,2006; Tanaka et al. 2008,2012) in column 4 of Table 1. Interestingly, it was not given for other studies with similar group sizes (and the same criticism could be stated for many of the studies used by FDA to develop the ADI). The concern about sample size seems to be specific to the dose-response step of quantitative risk assessment, but not to the value of the study as reviewed by OEHHA. Statistically significant effects detected with small samples are valid rejections of the null hypothesis and can identify hazard. Increasing sample size will always produce a smaller p-value for a given effect size whether the sample is large or small.

Comment 2.2: "…the study did not provide strong or consistent evidence of effects."

This comment was provided for all six Tanaka studies in column 4 of Table 1

Very similar wording was used in column 4 of Table 1 for the basis of their evaluation in all six Tanaka studies.

Tanaka (1994): "The results of Tanaka (1994) provide evidence of isolated changes that are not sustained over time despite chronic exposure to the test color"

Tanaka (1996): "of isolated changes that are not sustained over time despite chronic exposure to the test color"; "limited to one or two endpoints measured from a large number intended to evaluate neurobehavioral domains";

Tanaka (2001): "changes in a single endpoint for a battery of endpoints to assess exploratory behavior and observed at a single time point, but not at other timepoints, despite continued exposure to the test color."

Tanaka (2006): "changes reported in Tanaka (2006) are limited to single endpoints in a single sex at a single time point in the study"

Tanaka et al. (2008): "The changes reported in Tanaka et al. (2008) are limited to single endpoints in a single sex at a single time point in the study;"

Tanaka et al. (2012): "The results of Tanaka (2012) provide only limited evidence of isolated changes that are not sustained over time despite chronic exposure to the test color and are limited to one or two endpoints measured from a large number intended to evaluate neurobehavioral domains."

Response 2.2: Presumably "strength" refers to the "statistical power of results" as described in comment 2.2.1.1. Consistency is an adjective with a broad and subjective range of interpretation. OEHHA is not aware of recognized guidance for use of this term in toxicological risk assessment. Ramboll did not detail their criteria for "consistent effects".

In the absence of other criteria, this wording implies that effects documented in these studies would be expected by Ramboll to occur in both sexes, for all endpoints in all assays used, and at all timepoints included in the design. No specific endpoints and timepoints are detailed as inconsistent. The OEHHA review provides a table (Table 3.3) with all the behavioral activity results from the Tanaka studies and details on the specific endpoint, timepoint assessed and sex demonstrating the effect. Perhaps this table was used by Ramboll to identify "inconsistencies".

1. Consistency across sexes. Sexual differentiation of activity is widely recognized in rodents and was specifically documented in detail by the Tokyo laboratory using their rodent model and test apparatus (Tanaka 2015a). Sexual differentiation reflects differences in brain regulation of these behaviors and different baselines for detecting treatment effects, which does not support a presumption of consistency in treatment effects across sexes.
2. Consistency across endpoints. Activity testing yields a number of different endpoints designed to further specify the underlying biological processes responsible for the effect (Sobin and Golub 2018). For example, the parameters measured in the Tokyo laboratory (1994-2008) were number of movements, movement time, number of horizontal activities, total distance, number of vertical activities, vertical time, number of turns, average distance, average speed, and defecations. Vertical activity (rearing onto the back legs) could be affected by disrupted balance which would have little influence on horizontal activity (ambulation on the floor of the apparatus). Consistency between effects on these two endpoints would not be expected. Similarly, consistent findings across other combinations of neurobehavioral endpoints would often not be expected based on the underlying biology.
3. Consistency across timepoints. The Tanaka studies assessed activity at two timepoints, immediately after weaning at three weeks of age, and in adulthood at eight weeks of age. Sexual maturation and extensive brain development occur between these two ages providing a different substrate for biological actions of dyes. Further, the duration of exposure to the dyes is much longer at the older age, an important consideration if the dye produces progressive pathology. These considerations limit an expectation of consistency.

Two other studies are described in the Table 1 as inconsistent (Rafati et al. 2017, Sobotka et al. 1977). While Rafati et al. was described as not providing "strong or consistent evidence of effects", no further information was provided about the inconsistencies. Instead, the age of the rats was noted as a detrimental factor. For Sobotka et al., Ramboll stated "study did not provide strong or consistent evidence of effects" "a small transient improvement in development of neuromotor clinging ability in female neonates (which was defined as speculative)". Strength or consistency as applied to this effect is not described.

Notably, none of the mixture studies were mentioned by Ramboll for inconsistency in the Table 1 Column 4, in the Ramboll text, or in the IACM summary. In fact, a single identical comment

is provided by Ramboll for all the mixture studies and apparently no further evaluation was made. Mixture studies were dismissed from consideration for quantitative risk assessment by Ramboll, and from hazard identification by IACM, because effects could not be attributed to individual dyes. OEHHA agrees that the mixture studies cannot be used to attribute the effect to one or more of the food dyes. However, the OEHHA review emphasized the importance of the mixture studies because they were intended to parallel the studies in children, most which used mixtures. The relative amount of different dyes was selected based on human exposure. In reviewing the effects of toxic exposures, OEHHA is acutely aware that, while exposures are often assessed one at a time in the scientific literature, in real life exposures never occur in isolation. One attempt to address the issue of simultaneous exposures is to study mixtures using guidelines from USEPA (USEPA 1986, 2000).

Comment 2.3: Consistency of control group measures across Tanaka studies.

Also related to consistency in the Tanaka studies is the following comment included in column 4 of Table 1 for four of the six Tanaka studies (not clear why the other two were excluded):

"the results of the comparison of control groups indicate the model systems used by the Tokyo Metropolitan Laboratory of Public Health have significant variability over time, hindering the ability to draw any meaningful conclusions from the reported results."
Elsewhere (Appendix A, page 4) this comparison of control groups is attributed to Gentry et al (submitted)

"Gentry et al. (submitted) conducted a statistical evaluation to investigate the potential variability in responses in control groups over time. Comparable control data for comparison were reported across the Tanaka publications and included Tanaka (2001). Overall, the results of the comparison of control groups indicate the model systems used by the Tokyo Metropolitan Laboratory of Public Health (Tanaka 2001) have significant variability over time for the majority of endpoints reported to be statistically significantly different from controls."

Response 2.3: IACM did not provide this statistical evaluation and, as mentioned previously, Gentry et al. (submitted) is not publicly available. No further information was found by OEHHA in the IACM comments concerning which Tanaka studies were used, which descriptive statistics were compared, which analyses were used to evaluate variability, and most importantly, which endpoints were assessed for control group variability. Our own reviews did not find that quantitative data for neurobehavioral endpoints were reported similarly across Tanaka dye studies in a manner that would allow statistical analysis without raw data. (OEHHA attempted to contact Tanaka to obtain raw data but was not successful.)

It also seems that Gentry et al. may have tried to re-evaluate the statistics for some of the effects from the Tanaka studies using control groups combined over studies rather than the contemporary control group, based on this comment:

> "Further, combination of control results for comparison across treatment groups resulted in the loss of statistical significance in selected responses, because of the significant variability in control responses over time, hindering the ability to draw any meaningful conclusions from the reported observations upon exposure to any of the colors evaluated." (IACM letter, page 11)

Putting aside the lack of ability to examine the Ramboll comparison of control groups, OEHHA would like to emphasize three points:

1. A contemporary control group is required in toxicology studies as the basis of identifying treatment effects because uncontrollable environmental changes can occur. For example, the Tokyo lab has documented the effect of a change in bedding on their activity and maze learning measures (Tanaka 2014). Advancing animal husbandry sometimes requires changes in recommended environmental conditions.
2. During the 28 years the dye studies were conducted, the laboratory used a similar design to test the food dyes as well as a variety of other potential toxicants. A number of sequential protocol changes were made over the years particularly in the activity test which produced most of the significant dye treatment effects.

a. The duration of the activity tests changed. Tanaka 1994/1995-Tanaka 2001 and later.
b. The automated activity test apparatus was changed. Tanaka 2008-Tanaka 2012. The Tokyo lab documented in detail the differences between activity as measured in two apparatuses used sequentially over the years (Tanaka 2015b).

The statistical approach to evaluating treatment effects depends on the distributions of data in the individual experiment. Comparison of control groups across studies disengages them from the appropriate statistical framework for evaluating treatment effects in a given experiment. The Tokyo lab always referred to state-of-the-art approaches in selecting its statistical testing. A variety of tests were used over the years for the activity data: 1994/1996 Mann-Whitney test, 2001 Shirley-Williams test; 2006/2008 Steel-Dwass test; 2012 parametric statistics (ANOVA with Bonferroni group comparisons). Any re-calculation of the effects would require the raw data and include a preliminary evaluation of distributions for selection of an appropriate test.

Variability across studies in activity control group measures are likely attributable to such changes across the years and attest to the value of the contemporary control group.

Comment 2.4: A final issue related to consistency is consistency between studies for individual dyes as reflected in the following comments.

Noorafshan et al. 2018; "The findings are also questionable due to …… inconsistency with results from multi-generation studies in animals (Tanaka 1994)." (Table 1, column 4)

Dalal and Poddar 2009; . The findings are also questionable due to the inconsistency with results from multi-generation studies in animals (Tanaka 2001)." (Table 1, column 4)

Response 2.4: OEHHA began its review by pointing out:

"The exposure periods and assessment timepoints for the studies reviewed in this section are shown in Figure 3.2. There is great variety in these aspects of study design providing a depth of information, but precluding comparisons of findings."

The sets of two studies compared by Ramboll differed in animal model, route, doses, duration, and endpoints. This precludes identification of results for comparison to determine consistency.

OEHHA did provide extensive comparison of studies when appropriate including the two Tanaka studies examining Yellow No. 5 (Tanaka 2006, Tanaka et al. 2008), the two Dalal and Poddar studies examining Red No. 3 (Dalal and Poddar 2009, 2010), the two studies reported together by Vorhees et al. (1983a), and two mixture studies which were intended as replicates

(Shaywitz et al. 1979, Reisen and Rothblat 1986). These pairs of studies used similar designs.

The Ramboll statements on inconsistency are carried forward into the Ramboll summary and the IACM summary:
"All statistically significant outcomes appeared to be incorporated into the OEHHA summary of animal toxicology without evaluation or discussion regarding the lack of consistency of effects across studies, between sexes or life stages within the same study, or of the observed animal-to-animal variation in effects upon exposure to colors versus those in the control groups." (IACM comment letter, page 10)

"Based on Gentry et al. (submitted), none of the studies evaluated provided strong or consistent evidence of effects." (Appendix A, page 6)

In summary, while inconsistency is very frequently repeated as a reason to dismiss the OEHHA review by Ramboll and IACM, instances of inconsistency described appear to be based on application of an undefined criterion.

2.1.2. Response to Ramboll comments in Appendix A Table 1 column headed "Ramboll strengths/limitations of study"

OEHHA notes that comments in this column were sometimes carried over to column 4. They often reflect the items in the ToxRTool score sheet.

Comment 2.5: "Study relied upon the mouse as the animal model instead of the rat, the rat being the recommended species in the OECD/ICH test guidelines."

Response 2.5: This comment was given for all the Tanaka mouse studies. Tanaka et al. (2008,2012) state that their study is based on OECD guidelines adapted for mice. This is the only OECD recommendation that was mentioned by Ramboll in the Table 1 as a criterion for limitations.

Comment 2.6: Four studies were excluded by Ramboll from consideration "in any quantitative evaluation of neurobehavioral effects in children" because of age:

Noorafshan et al. 2018: "the study was conducted in mature adult rats"

Dalal and Poddar 2009: "study was conducted in young adult rats".

Dalal and Poddar 2010: "study was conducted in young adult rats"

Ramboll noted that two other studies were confounded or uncertain due to age of animals when exposed or tested:

Rafati et al. 2017: "advanced age is a confounder".

Erickson et al. 2014: "Significant uncertainty would arise because exposure was only during a limited time during juvenile development".

Response 2.6: Neurobehavioral studies in adult animals were in line with OEHHA's goal of examining food dye neurotoxicity and were reviewed in Section 3.3 "Adolescent/adult

Neurobehavioral Toxicity Studies" which concluded with Section 3.3.7 "Relevance of adult neurotoxicity studies to developmental neurotoxicity issues." A portion of that section said:

"Lifestage specificity of neurotoxicants: The issue of food dye effects on children falls outside of the regulatory framework for developmental neurotoxicity, which focuses on long-term or lasting effects of an exposure limited to development. Some types of developmental neurobehavioral toxicity, however, can be seen to parallel adult neurotoxicity where dosing and testing are done at the same time."

We do not know whether neurobehavioral toxicity of food dyes in humans is specific to a lifestage that excludes adults as no research is available on that topic. Further, adolescence has not been specifically addressed in human food dye studies.

The comment regarding "advanced age" in Rafati et al. may have been a misreading by Ramboll. The paper (Rafati et al. 2017) describes the animal model as "seventy two-month old male Sprague-Dawley rats". Further reading of the methods section clarifies that they meant seventy rats who were two months of age. Ramboll may have read this instead as rats who were seventy-two months old. In any event, age was not a confounder as the rats in all groups were all the same age when they started the experiment and when they were tested in the radial arm maze.

Regarding the comment on Erikson et al., dosing in this study was conducted from 22 to 50 days of age. This covers the entire period from weaning to sexual maturation. This juvenile development period in rats is parallel to childhood in humans and thus is appropriate to evaluate and especially relevant for comparison to food dye studies in children.

Comment 2.7: Food restriction. Noorafshan et al. 2018: Ramboll notes that "In addition, the Radial Arm Maze test for learning and memory was conducted under altered feeding regimen to force a body weight reduction of 15%. It is unknown what the effect of the forced-food restriction could have had on the bioavailability of the test substance or the results of the learning and memory test in general." (Table 1, column 4)

This comment is further amplified in the IACM summary (page 12):

"The authors offered no details on the potential effects of the forced food restriction and reduction of body weight on the bioavailability of the test substance, the effect of 15% body weight reduction on the overall outcome compared to normal weight control animals, or the relevancy of the forced food restriction to normal human behavior."

Response 2.7: All maze tasks use reinforcement to indicate the correct response for learning and memory. As cited by Noorafsham (Pritchett and Mulder 2003), and confirmed in other publications (Wenk 2004), food restriction Is part of the protocol for food-reinforced radial arm maze tests. Maintenance of body weight at 85% of baseline is recommended as a regimen that provides appropriate motivation without ill effects. Since the dye was administered by gavage, the food restriction does not affect dose. A potential effect of food restriction on bioavailability from gavage is not described by Ramboll. Human learning and memory tests use motivations appropriate for humans.

Rafati et al. (2017) also used a food reinforced radial arm maze. While food restriction is not mentioned in that article, it is described in the citation used as the source of the procedure (Karkada et al. 2012).

Mazes can also be reinforced by escape from water, which was used as the reinforcer for the maze test in the Tanaka studies. It is also possible that immersion in water as well as food restriction could affect test agent ADME. The most recent guidelines for evaluating developmental neurotoxicity studies discuss food-reinforced learning and memory maze tests recognizing the role of food restriction in motivation (NAFTA 2016, p.D-9).

Comment 2.8: Washout period. Gao et al. 2011, Rafati et al. 2017: Ramboll notes these investigators "Introduce uncertainty into the analysis by including a "wash-out" period before recording behavioral, molecular, and morphometric measurements on different PNDs, impeding the interpretation of the results in the context of meaningful toxicological outcomes." (Appendix A, Table 1, column 2 – Ramboll strengths and limitations)

Response 2.8: Ramboll does not say what a "wash-out" period is, but it might mean a lag between the end of dosing and the beginning of measurement. OEHHA was unable to find any indication of a "wash-out" period in these studies. In the Rafati et al. (2017) study dosing was administered for seven weeks, behavior testing started after four weeks and animals were sacrificed for brain measures at the end of dosing. In the Gao et al. (2011) study dosing of mice was administered for 30 days and testing started the day after dosing started. The rats were also dosed for 30 days and sacrificed after completion of behavior testing for brain measures.

Comment 2.9: Gavage vs. diet. In Appendix A, Table 1, column 2, Ramboll lists gavage administration as a study quality "limitation" for seven studies (Noorafshan et al. 2018; Dalal and Poddar,2009,2010; Gao et al. 2011; Rafati et al. 2017, Shaywitz et al. 1979, Reisen and Rothblat 1986). Administration in diet is given as a study strength.

Response 2.9: It is not clear why Ramboll considered gavage administration a limitation. Elsewhere IACM states that exposure was "via oral gavage and not diet, which would have emulated human exposure". In its review, OEHHA emphasizes the value of animal studies with gavage dosing because "They speak to the ability of food dyes to affect behavior when the entire daily dose is administered at one time as is the case in children's challenge studies." Note that a common everyday source of food dyes is children's drinks. Drinks, which are in liquid form and not embedded in a food matrix, can be consumed fairly rapidly and as such are more like gavage than dietary administration.

Comment 2.10: Dosing. "Only two concentrations plus negative control tested." Ramboll cited this limitation in Table 1, column 2, for all studies with fewer than 3 doses.

Vorhees et al.1983b: "The remaining endpoints assessed were sporadic, with a lack of dose-dependence reported."

Vorhees et al.1983a: "the non-dose- dependent nature of these effects, combined with their lack of replicability across experiments, renders these findings unconvincing as evidence"

"Doses in the study were magnitudes higher than the ADIs and higher than expected in the human population." This Ramboll limitation was cited for all the Tanaka studies in Table 1, column 2.

Response 2.10: Limitations concerning the number of doses, and the dose-dependence of findings may be related to Ramboll's focus on excluding studies not suitable for quantitative risk assessment. OEHHA's review discussed dose trends for each multiple dose study with an eye to the increasing recognition of nonlinear dose-response relationships as biologically relevant.

The basis for the comment regarding doses magnitudes higher than the ADIs is not clear. Toxicology guideline studies include a maximally tolerated dose. The Tanaka studies, by using doses 100 or 10 X the ADI, would be providing information for doses that could confirm or modify the ADI because the 100 X safety factor is typically applied to animal studies when deriving ADIs.

Comment 2.11: Beneficial vs. adverse effect

2.11a. Tanaka (2006), Tanaka et al. (2008): Ramboll notes "direction of accelerated achievement of coordination indicating better performance in treated versus control animals". (Table 1, column 4)

Response 2.11a: OEHHA did not place value on the developmental test batteries in the Tanaka studies as they did not use litter-based statistics. Further, in general, any deviation from normal patterns of development needs to be carefully considered as potentially an adverse effect.

2.11b. Tanaka (2001): Ramboll notes "Results from the exploratory behavior test were inconsistent and in some cases the changes could indicate an improvement in exploratory behavior and not an adverse treatment-related effect". (Table 1, column 4)

Response 2.11b: Ramboll does not provide any detail on which tests were inconsistent or indicated an improvement. In general, deviations from control group patterns of exploration can be taken as indication of disruption of regulation of this spontaneous behavior. Over-exploration is a sign of disinhibition, which is adverse.

2.11c. Regarding Vorhees et al. (1983b) study on Red No. 40, Ramboll notes "The authors acknowledge the inconsistency in endpoints in that one is associated with hypoactivity and the other hyperactivity". (Table 1, column 4)

Response 2.11c: The comment misreads the Vorhees et al, 1983b paper, which indicates clear effects on behavior in a reproductive and developmental study where rats were given Red No. 40 in the diet. The full statement of the authors is as follows:

"The consistent behavioral effects, taken together with the other measures of toxicity, indicate that R40 at the doses used here is both physically and behaviorally toxic, and that the manifestation of the behavioral toxicity was most evident on measures of activity. The present data do not support, however, a concept that psychotoxicity from food colors will be reliably manifest only or even primarily as hyperactivity, since in the present experiment, hyperactivity was found in one test of activity (the open-field), while hypoactivity was found on a second test of activity (the running wheel)."

OEHHA notes that the running wheel test measured wheel turns 24 h/day over a 21-day period in the home cage, while the open field test measured ambulation for 3 minutes for three consecutive days in a novel environment. While activity was the general domain measured, the tests measured different aspects of activity and were not designed to generate consistent effects (e.g., both endpoints going in the same direction of increase or decrease) in that domain.

2.11d. Dalal and Poddar (2010): Ramboll notes "Further the changes were a decrease in activity in contract (sic) to concerns in children regarding hyperactivity." (Table 1, column 4)

Response 2.11.d: In contradiction to the Ramboll statement, the Dalal and Poddar (2010) study showed an increase in activity after dye administration (OEHHA, Figure 3.4). Ramboll might have meant Dalal and Poddar (2009) which showed a decrease. The experimental designs distinguished between these two findings and are discussed in detail in the OEHHA review on pages 98-102. In essence, dosing with dye resulted in a change in behavior in this rodent model, which correlated to serotonergic changes in the brain. The concept that a single administration resulted in decreased activity while administration after 15 or 30 days exposure resulted in increased activity is discussed in the OEHHA report. As we state on page 102, "The authors [Dalal and Poddar] attribute the differing effects of Red No. 3 with and without prior daily exposure to elevated corticosterone status with repeated exposures which increases brain serotonin synthesis and synaptic levels causing increased activity."

Comment 2.12: Identification of NOAEL/LOAEL, Vorhees et al. (1983a) "While OEHHA concluded the lowest dose tested was the LOAEL (a dose higher than the FDA ADI), this is inconsistent with the authors conclusions for behavioral effects. (Table 1, column 4)

Response 2.12: The full statement of the authors is as follows:

"Tests of behavioral development and adult performance showed several statistically significant changes, but the non-dose-dependent nature of these effects, combined with their lack of replicability across experiments, renders these findings unconvincing as evidence of psychotoxicity."

OEHHA was able to use the statistically significant effects to identify a LOAEL. Our review did not consider the lack of linear dose-effect relationships as reason to disregard the effects. There are a number of examples where the dose-response for a toxicant is non-linear (e.g., supralinear dose-response at low doses of lead; u shaped dose-response for a number of endocrine disrupting compounds). Also, OEHHA examined replicability in detail and did not find it restricted identification of a LOAEL because true replication was not achieved:

"However, the two experiments differed extensively, particularly in the general level of preweaning mortality, and the increased statistical power of the ANOVAs in the first experiment due to inclusion of two control groups."

Comment 2.13: When studies were conducted. Randoll notes "study conducted (*34,36,37,40,41*) years ago"

This Ramboll limitation was used for five studies conducted more than 33 years ago.

Response 2.13: OEHHA reviewed all studies in the context of the era in which they were produced and attempted to provide continuity in hypothesis testing across the years, a hallmark of investigator-initiated research. We used expert judgment of the validity of all the reviewed studies (see OEHHA Table 3.9 and 3.10). We used a broader set of criteria in evaluating studies than the criteria in tools such as SciRAP or ToxR. In terms of the value of these studies, it may be noted that all the FDA ADIs were derived more than 33 years ago and thus were based on studies conducted prior to that time.

Comment 2.14: Interpretation of nonbehavioral measures

Ceyhan et al. 2013: Ramboll notes "toxicological meaning of these changes not completely understood". (Table 1, column 4)

Dalal and Poddar (2009): Ramboll notes "The study consistent of a single oral gavage dose, with transient changes in neurotransmitter levels reported with no associated change in brain morphology." (Table 1, column 4)

Response 2.14: An exciting development in recent investigator-initiated food dye research is examination of brain tissue in studies that show behavioral effects. In these research papers, the investigators provide a rationale for the measures they selected and an interpretation of the results. Ramboll's criticism that they would not include these studies because the effects were transient and there was no change in brain morphology does not reflect the potential for harm from a transient effect that is repeated several times every day. In fact, a transient effect on neurotransmitters after a single dose of dye is relevant to the time dependent change in behavior of children after a dye challenge as documented in the children's studies. The Dalal and Poddar study did not measure brain morphology, thus we don't know whether or not there were changes in brain morphology.

2.1.3. Comments from Ramboll text not based on Ramboll tables.

Comment 2.15: Use of study evaluation tools. Ramboll states "While a study quality assessment was conducted for the human studies, the animal toxicity studies that were considered in the OEHHA (2020) Draft Report were not evaluated by OEHHA using any type of published methods for study quality or reliability." (Appendix A, page 1)

Response 2.15: The Ramboll assessment of study quality is referred to in Gentry et al. (submitted). Ratings from this assessment used ToxRTool and ARRIVE and are included in the Appendix A Table 1 and repeated throughout the Ramboll text and the IACM summary text. Ratings from SciRAP are provided in Appendix B.

We are not able evaluate this quality review because Gentry et al. (submitted) is not publicly available. The utility of the ToxRTool has been evaluated by a panel of 8 scientists from various institutions who identified strengths and weaknesses and made recommendations for improvement (Segal et al. 2015). A conclusion was that the ratings are primarily based on study reporting quality, noting that reporting quality and study quality are not the same. The ARRIVE guidelines are specifically for study reporting. OEHHA did not use a summary tool as a format for providing information on study reporting but included the information in categories in the summary tables (Tables 3.9 and 3.10). In addition to study reporting, OEHHA evaluated the quality of all aspects of the research and provided detailed strengths and weaknesses. It may be worth noting that as a result of Ramboll use of ToxRTool only three studies that OEHHA reviewed were judged "not reliable". The rest of the studies were rated by Ramboll as reliable without restrictions or reliable with restrictions.

Comment 2.16: (Appendix A, page 1) Quantitative Risk Assessment

Scope of studies reviewed. "The studies that initially appeared to be the focus of OEHHA's quantitative risk assessment (QRA) for individual colorants, based on the 2019 symposium presentation, were the studies conducted predominantly in the Tokyo Metropolitan Laboratory of Public Health (Tanaka 1994, 1996, 2001, 2006; Tanaka et al. 2008, 2012) in addition to a few others (Doguc et al. 2013, 2015, 2019; Ceyhan et al. 2013; Başak et al. 2014, 2017). An additional 13 animal studies were considered for quantitative risk assessment in OEHHA's 2020 draft report. OEHHA (2020) states that primarily developmental studies with oral administration and neurobehavioral endpoints were included in the review; however, OEHHA (2020) also included studies when dye exposure was only during adulthood in the animals.

These studies do not address the charge. In addition, if non-behavioral endpoints were included in the study, they were also reviewed. Therefore, the studies under consideration were broader than the original OEHHA presentation implied."

Response 2.16: The comment appears to imply that we should have stuck to the few studies we began with during our initial assessment of the literature prior to the workshop we held in October, 2019, to solicit early input. Rather than commenting on OEHHA's review, Ramboll presents a parallel review based on considering or excluding studies for quantitative risk assessment. As detailed above (response 2.1 above) OEHHA did not conduct a traditional quantitative risk assessment. As detailed in the Introduction to the OEHHA report, we conducted a hazard identification and risk characterization using a variety of approaches (see OEHHA report chapter 7). The breadth of studies reviewed reflected the mandate to conduct a scientific review of the potential for neurobehavioral effects of synthetic food dyes in children. The parameters for the literature review are included in the OEHHA document. OEHHA read and integrated information across many disciplines as is appropriate for scientific review. Note that studies conducted in adults provide information on the capability of the synthetic food dyes to cause behavioral change or other neurological endpoints – we would be remiss if we did not include these studies.

Comment 2.17: (Appendix A, page 2) Quantitative Risk Assessment. "In reviewing the Tanaka studies, Ramboll (2020) noted that the comparison of the results across all of the Tokyo Metropolitan Laboratory of Public Health studies can be conducted to evaluate the implication for a quantitative risk assessment, as the chemical composition ofthe colors would suggest similar results might be expected."

Response 2.17: OEHHA notes that Ramboll (2020) is not publicly available and cannot be evaluated, but, in general, it is recognized that the six dyes belong to structurally distinct chemical groups with distinct biological and toxicological properties. This is expressed in the IACM summary text:

"The allowed (certified) color additives in the U.S. include azo dyes, a triarylmethane derivative, an indigotine derivative, and an iodofluorescein derivative. These chemical classes are significantly distinct in chemical structure. Therefore, ascribing behavioral effects to all FD&C colors via a unified mechanism is not scientifically supported, is not addressed nor demonstrated in OEHHA's draft report, and should be carefully questioned."

Note that OEHHA does not ascribe all the same effects to all of the different synthetic food dyes. We evaluated published animal studies of individual food dyes as well as mixtures of dyes. We describe reported effects for each of the dyes as observed by a variety of investigators.

Comment 2.18: (IACM comment letter, page 11) "Also, while OEHHA notes that the JECFA ADI for Red No. 3 is two orders of magnitude lower than the FDA ADI, it is worth noting that it is actually only one order of magnitude lower."

Response 2.18: Thank you for pointing out this inaccuracy. This has been corrected in the text. The point of departure for the Red No. 3 JECFA ADI was two orders of magnitude lower than the point of the departure for the FDA ADI. Different safety factors led to the one order of magnitude ADI difference.

Comment 2.19: (IACM comment letter, page 12) Mechanistic data. "While the OEHHA report mentions various *in vivo* and *in vitro* studies of mechanistic data for each color, there is no overall conclusion regarding how such data contribute to the understanding of underlying mechanisms of neurobehavioral effects, nor are such data considered in the context of plausible modes of action or adverse outcome pathways related to neurobehavioral outcomes. There is an overall lack of commentary on dose (relevance to human exposure), duration/timepoints, models used (predictability or consistency of models used, either *in vivo* or *in vitro*), or "bridging" of molecular or cellular signals to behavioral effects in the OEHHA report. It is important to note that a single signaling event is not enough for an adverse outcome pathway to be relevant, but rather multiple key events (KEs) are required to get from the molecular initiating event (MIE) to the adverse outcome"

Response 2.19: At this point in the research on the mechanisms of neurobehavioral effects of the food dyes, hypothesis-testing studies are being conducted; more are needed. The OEHHA review tries to identify possible mechanisms by which the observed effects in animals and children might occur. As well, we tried to identify potentially fruitful avenues of investigation. Adverse outcome pathways (AOPs) provide a means of identifying potential mechanisms of toxicological effects. AOPs do not necessarily verify a sole pathway by which an adverse effect to a toxicant can occur; there may be multiple pathways to an outcome. Further, for most chemicals and toxicological endpoints, there are not enough data to identify AOPs, molecular initiating events or key events. Finally, AOPs do not limit scientific inquiry concerning mechanisms of toxicological effects that have already been identified in human and animal *in vivo* studies.

3 Response to comments in Appendix B on mechanism studies from the peer-reviewed literature.

Introduction to OEHHA responses

Appendix B of the IACM comments was provided by ToxStrategies (TS). This appendix compares the literature search that TS conducted to OEHHA's literature search, citing Gentry et al. (submitted) and Chappell et al. (2020). The appendix provides general comments, which we address first below, and comments on individual dyes.

Comment 3.1 (**Appendix B, page 2)** Study reliability. "Study reliability (i.e., quality of methods and reporting) was not considered in the OEHHA assessment. While study reliability was not used as a means to include or exclude studies in the mechanistic investigation reported elsewhere,[3]/ study reliability was evaluated for each study using the SciRAP tool."

The IACM Appendix A also mentions the SciRAP tool (Appendix A, page 1):

"many of these studies were assessed independently using the Science in Risk Assessment and Policy (SciRAP) tool"

Response 3.1: The comment indicates that TS used the Science in Risk Assessment and Policy (SciRAP) tool when they conducted their own literature search. Results of the application of the SciRAP tool are mentioned throughout Appendix B and attributed to Gentry et al. (submitted). The text in this comment implies that lack of use of summary tools is the same as lack of consideration of quality. OEHHA considerations of quality and reliability

issues are included in the text of the review. In addition, in Tables 3.9 and 3.10 we describe some of the factors that we considered in our study quality review, which are similar to those described in the ToxR Tool.

Notably, the SciRAP tool is not mentioned in Chappell et al. (2020) paper or Supplementary materials. OEHHA is unable to review the validity of the SciRAP tool results because Gentry et al. (submitted) is not publicly available and was not provided to OEHHA by TS**.** Citations to Gentry et al. (submitted) by TS sometimes include specific subdivisions, for example, **"**detailed in Table S8-1 in Supplement 8 in Gentry et al. (submitted)[3]". It is not possible for OEHHA to respond without the cited material.

Comment 3.2: (Appendix B, page 2) Mechanism study search strategy. "The OEHHA assessment used a broad search strategy to identify studies associated with each color and did not seem to use filters with a clear objective for sorting the animal and/or *in vitro* information relevant to neurobehavioral endpoints. As such, they identified and included studies in their evaluation that others excluded or, in several instances, did not identify in the literature search.[2/3]". The footnote ([2/3]) identified "others" as Chappell et al.(2020) and Gentry et al.(submitted). The Chappell et al. paper is from ToxStrategies and the Gentry et al. (submitted) paper is apparently a collaboration between Ramboll and ToxStrategies.

Response 3.2: This comment generally confirms that OEHHA used a scientific review process different from Chappell et al. (2020) (Since Gentry et al. was not available to OEHHA at the time of responding to these comments, we did not have much information on their review process).

There were notable differences in review process between Chappell et al. and OEHHA:

- Chapell et al. conducted a literature search, identified 83 studies based on key words, eliminated 71 based on review of titles and abstracts, and included 12 in their review.
- OEHHA used different keywords for the literature search and did not eliminate any studies from review before reading them. After reading them, we included the ones that were relevant, reading further on the topic if necessary. According to Appendix B, OEHHA reviewed 79 mechanism studies.
- Chapell et al. used "the inclusion criterion that any mechanistic data should map to neuro-relevant pathways".
- OEHHA included any studies that shed light on dye biological actions relevant to behavioral effects and risk characterization.

The TS review of Yellow No. 5 literature provides an example of the differences between OEHHA's review and that of Chappell et al.:

- Chappell et al. (2020) retrieved 10 articles from their literature search and identified none as being relevant.
- OEHHA Identified 24 studies to review.
- In Appendix B, TS apparently found 5 studies from the OEHHA review as relevant which were not found by Chappell et al. (2020).
- The other 19 studies reviewed by OEHHA were excluded as being "out of scope", "not neurorelevant", or "evaluation of mixtures".

Our effort to identify as much relevant information as possible was clearly broader than the scope of the Chappell et al. (2020) review. Our broader scope was designed to find any and all information that would inform whether the food dyes potentially impact neurodevelopment. For

example, TS eliminated all mixture studies because TS was focused on dose-response assessment. As well, the comment does not define what they meant by "neurorelevant" or how they decided which studies were "neurorelevant".

Comment 3.3: (Appendix B, page 2) A third comment has to do with protein binding of Red No. 3.

"For *in vitro* studies, the significance of the extent of protein binding on the evaluation of neuro-relevant mechanistic data, especially *in vitro* evaluations, is unclear (e.g., Red No. 3)";

"…although change in neurotransmitter function remains a key event that could in combination with other key events potentially lead to neurobehavioral adverse outcomes, many of the studies (especially the *in vitro* studies) did not consider artifacts at the time these studies were conducted that would otherwise potentially erroneously produce the assay results noted. These artifacts include non-specific protein binding characteristics of Red No. 3".

Response 3.3: As mentioned by both OEHHA and TS, there is a large literature on Red No. 3 and neurotransmitter release. There is also a literature on Red No. 3 protein binding. An historical overview of the intersection of these two areas of research can help clarify "the significance of the extent of protein binding". This perspective is in the OEHHA review but provided again here.

After several studies were published in the 1970's on Red No. 3 and neurotransmitter uptake and release, Mailman et al. published a study in Science (Mailman 1980). This is the only peer-reviewed data published by Mailman on the topic of dye protein binding. A SciRAP score for this study is not provided by TS but the methods were dependent on the era and mainly descriptive. There is no information on the source or purity of the Red No. 3, the duration and temperature of the incubation or other report quality criteria used in SciRAP. Protein binding was inferred from the amount of the tissue preparation in the incubation. The amount of protein at two tissue concentrations was determined by an unidentified assay. While the amount of protein varied with the tissue concentration, so did all the other components of the extracted tissue. A graph was presented showing the uptake of one concentration of labelled dopamine at five concentrations of tissue. (Behavioral data were also presented in the article but were not reviewed by OEHHA due to the lack of methodological detail.) No replication or follow-up studies were published. Other citations of this author by Chappell et al. (2020) and TS (Mailman 1987, Mailman and Lewis 1987) are reviews published later in the 1980s. In addition to the Mailman study, Red No. 3 protein binding was inferred to be an important factor in distribution to brain from systemic circulation (Levitan et al. 1984,1985) as reviewed by OEHHA. However, prior to 2000, little information was available on protein binding as it relates to Red No. 3 toxicology.

The extensive protein binding of food dyes is an important factor in providing uniformity and stability of coloring of food. Also, many food dyes we reviewed have been used for years in laboratory protein identification assays. Because of this, dyes we reviewed are often used as model agents in contemporary basic research on protein binding.

Currently, protein binding studies are much enhanced by mathematical modeling that identifies binding sites and the binding potentials for individual sites. The importance of binding to protein aggregation in neurodegenerative disorders is under study. This literature needs to be experimentally connected to food dye toxicology so we included citations in the OEHHA review. Red No. 3 protein binding studies cited by OEHHA (Ganesan et al., 2011; Ganesan

and Buchwald, 2013; Lee et al., 2016; Wong and Kwon, 2011), but not reviewed by TS due to inclusion criteria, are an important gateway to the protein binding literature that is needed to understand and enhance all food dye toxicology studies. At this point, just a few conclusions are appropriate:

- Protein binding is not an artifact in some studies. It is relevant to all *in vitro* assays, including the ToxCast assays. If the dye produces an effect *in vitro* in a targeted assay it is finding a target. The concern about nonspecific protein binding would be the extrapolation from the effective doses *in vitro* to human oral exposures.
- This issue is not specific to neurotoxicity. Any study of food dye toxicity mechanism has the same considerations. *In vivo* toxicity studies also need to think about the effects of protein binding on absorption, distribution and metabolism.
- Widespread protein binding is a positive effect for coloring foods evenly, but biologically widespread promiscuous protein binding could lead to many cellular changes, for example, protein conformational changes leading to inactivation of enzymes.

3.2 Response to study specific comments in Appendix B.

3.2.1. Studies not included in the OEHHA review.

TS mentioned a number of studies they reviewed that were not included in the OEHHA review. Presumably these studies were reviewed in Gentry et al. (submitted), as they were not cited in Chappell et al. (2020). Thus we have no idea why and how they were reviewed because Gentry et al. (submitted) is not publicly available and was not provided by IACM.

Comment 3.4: a. (Appendix B, page 7) "Six studies that provided mechanistic data were reviewed by TS but not by OEHHA (Bernstein et al., 1975; Butterworth et al., 1976a; Butterworth et al., 1976b; Capen and Martin, 1989; Hansen et al., 1973b; Hiasa et al., 1988)"

Response 3.4.a: These studies were in our bibliography and all concern Red No. 3 effects on thyroid-related endpoints. A summary review (Capen 1998) cited by OEHHA includes these studies. The context of the review is carcinogenesis because that was a major concern at the time the studies were conducted.

Comment 3.4.b. (Appendix B, page 9) "Six additional *in vivo* mechanistic studies conducted in experimental animals as reviewed by TS in Gentry et al. (*submitted*)[3]/ yet not included in the OEHHA assessment showed an overall equivocal response in various measures of thyroid pathway effects following oral exposure to Red No. 3, with effects most often observed at high dose levels (4% in diet) (Bernstein et al., 1975; Butterworth et al., 1976a, 1976b; Capen and Martin, 1989; Hansen et al., 1973b; Hiasa et al., 1988; detailed in Table S8-1 in Supplement 8 in Gentry et al. (submitted).

Response 3.4.b: TS provides this descriptive summary without any other information on study description, study quality or reliability, or description of their relevance to the OEHHA review.

Comment 3.4.c. (Appendix B, page 10) "The OEHHA assessment did not include two studies identified and reviewed by Chappell et al. (2020), which investigated the potential effects of Red No. 3 on central catecholamine systems and central nervous system injury and reported a lack of treatment effect (Mailman 1987; Mailman and Lewis, 1987)"

"Notably, none of the critical studies conducted by Mailman (1987) and Mailman and Lewis (1987), as cited by Chappell et al., 2020, were even considered by OEHHA (no references to these studies were made)"

Response 3.4.c: These two papers are reviews and do not provide any new data. Chapelle et al. (2020) apparently counted these reviews as research reports. After his 1980 publication in Science (Mailman et al. 1980), Mailman published a number of reviews over the years. An abstract and chapter authored by Mailman were also cited in Mailman's reviews. All the research on Red No. 3 that was reviewed in Mailman (1987) was included in the OEHHA review, except for one study of artificial membrane permeability. This study, Colombini and Wu (1981), was also not reviewed by TS.

Comment 3.4.d. (Appendix B, page 13) "Two mechanistic studies were reviewed by TS but not OEHHA (Albasher et al., 2020; Alsalman et al., 2019)"

Response 3.4.d: These two Yellow No. 5 studies were not available when the OEHHA document was prepared. Albasher et al 2020 did not use litter-based statistics, and the Alsalman paper focuses on carcinogenic effects. OEHHA did not consider the studies as important for addition to the document at this point.

3.2.2 Lack of OEHHA conformity with inclusion criteria of Chappell et al. and Gentry et al. (submitted) (Appendix B)

Comment 3.5: (page 17) "The other study reported by Goldenring et al. (1982) was designed to evaluate behavioral changes of rat pups following chronic injection (ip administration) of one dose level of sulfanilic acid, a metabolite of Yellow No. 6, with and without pre-treatment of 6-hydroxydopamine. As such, this study did not meet the inclusion criteria identified for evaluating neuro- relevant mechanistic data following oral administration to Yellow No. 6."

(page 17) "The OEHHA assessment (Section 4.1.4, page 142) stated, "This finding with sulfanilic acid suggests it is the active agent for Yellow No. 6 effects on cholinergic systems, as well as for effects on behavior (Goldenring et al., 1982) and identif[ies] this neurotransmitter system as a potential mechanistic pathway for Yellow No. 6 neurotoxicity." However, Osman et al. (2004) is of low quality/reliability based on the SciRAP assessment" and the lack of information on test-article purity, along with being conducted at only one exposure concentration,[3]/ thereby precluding any dose-response evaluation"

(page 17) "Four *in vitro* studies from the literature were discussed in the OEHHA assessment (OEHHA Section 4.1.3 and 4.1.4, pages 140–142), none of which were identified by Chappell et al. (2020).[2]/ One of these studies reported potential estrogenic effects (Axon et al., 2012) and was considered out-of-scope by Chappell et al. (2020), and another study was excluded as it was evaluating effects from a mixture of colors (Park et al., 2009)".

(page 2) "Another study (Chen et al., 2016) was not identified in the search conducted by Chappell et al. (2020) because studies on protein binding without measures of neuro-relevant mechanistic events were not included in the objective of that analysis." (Appendix B, page 2)

(page 4) "Seven studies identified by OEHHA (Section 4.1.6, page 144) provide toxicology information from oral exposure routes (Hollingsworth, 1982; Kobylewski and Jacobson, 2012;

Borzelleca and Hogan, 1985; Borzelleca et al., 1985; Butterworth et al., 1975; Gaunt et al., 1969; Hansen et al., 1966a). These studies were either not identified within searches conducted by others[2/3]/…… or were identified but excluded from the assessment because they did not provide neuro-related mechanistic data"

Response 3.5: These and similar comments were provided by TS concerning studies reviewed by OEHHA. TS pointed out that the studies did not meet the criteria of the Chappell et al. and Gentry et al. reviews. These criteria involved study quality and reliability criteria using category summarizing "tools", as well as relevance deduced from screening title and abstracts, and a definition of "neuro-relevant" that was not provided.

As mentioned above, OEHHA performed a broad, in-depth scientific review that did not conform to the evaluation format used by Chappell et al. (2020) and Gentry et al. (submitted). For some dyes (Blue 1, Blue 2, Green 3, Yellow 6) the minimal amount of literature garnered with neuro-related key words including "brain" motivated OEHHA reviewers to read broadly on the biological activity of the dyes. One function of a literature review is to bring together information from multiple streams that might not have intersected previously. We looked for biological actions in tissues other than brain that might generalize at the cellular level to nervous tissues and effects via routes other than oral that could be relevant to compromised or immature GI tracts. See Section 3.2.3.5. Using the SciRAP tool would not allow for the necessary broad review and search for information relevant to biological effects of the dyes. Given the relatively small database of studies, we reviewed everything, and thus did not have structured inclusion/exclusion criteria.

3.2.3. Response to specific comments on content of reviewed studies in Appendix B.

Comment 3.6: (page 2) regarding Lau et al. (2006): "Only three *in vitro* studies from the literature were discussed in the OEHHA assessment (Section 4.1.5, page 143), one of which (Lau et al., 2006) was also included in the assessment presented in Chappell et al. (2020). While both reviews noted the neurite outgrowth inhibition effect of Brilliant Blue in mouse neuroblastoma cells, the OEHHA assessment did not mention the cytotoxicity results reported by Lau et al. (2006), in which approximately 35% cell death was observed at a concentration over 1000x lower than the IC50 for neurite outgrowth reduction".

Response 3.6: This statement appears to be derived from Figure 6 in Lau et al. (2006). Figure 6 provides cell death data for the IC_{20} concentration used by the investigators in their mixture studies. The cell death data for .05nM is shown as a bar in Figure 6 that could be estimated as ~35%. This concentration (.05nM) is 1000x lower than the IC_{50} (51nM). The cell death value is from the individual dye and the media used to induce differentiation of the neurons. The "CSF" bar could represent media alone although this is not specifically stated.

The authors were concerned about a possible relationship between cytotoxicity and effects on neuronal differentiation of the mixtures. As they point out, other studies show that Blue No. 1 is a mitochondrial toxicant. They distinguished these two effects in their conclusions on mixtures:

"When assessing cell death mechanisms of food additive combinations, we found that both combinations studied had a straightforward additive effect on cell viability, as measured by

Trypan Blue dye exclusion. The mechanisms of synergistic neurotoxicity are therefore unrelated to effects on viability."

Comment 3.7: (page 17) " Although the *in vivo* study described by Osman et al. (2004) was of poor quality, it supports previous *in vitro* findings reported by the same investigators (Osman et al., 2002) in which Yellow No. 6 inhibited human cholinesterase and pseudocholinesterase activity, discussed by OEHHA (Section 4.1.4, page 142). However, OEHHA also stated that "In a second experiment, the IC50 for sulfanilic acid inhibition of cholinesterase and pseudocholinesterase was also demonstrated *in vitro* with a lower potency than Yellow No. 6." It is not clear what study OEHHA was referring to since the Osman et al. (2002) study did not evaluate sulfanilic acid."

Response 3.7: The reference should have been Osman et al. (2004). This has been corrected in the final document.

Comment 3.8: (page 2) purinergic receptors. "A study by Wang et al. (2013) that identified alterations to purinergic signaling was excluded from the Chappell et al. (2020) assessment based on the fact that purinergic receptors (specifically P2X7) are involved in Huntington's disease (which is inherited), spinal cord injury, "and other neurodegenerative diseases involving neuroinflammation" (as specified in Wang et al., 2013), as opposed to neurobehavioral effects or disorders related to children"

"OEHHA states (Section 5.3.5, page 169), "Both Blue No. 1 and Green No. 3 inhibit purinergic receptors"; however, the relevance of this finding to potential neurobehavioral outcomes is unclear because purinergic receptors are potentially related to hereditary neurodegenerative diseases involving neuroinflammation (e.g., Huntington's disease) as described for Blue No. 1 by Wang et al., 2013), or have been shown to have an antinociceptive (i.e., beneficial) effect (related to exposure to Green No. 3, as described in Yang et al., 2019). The relevance of purinergic signaling to neurobehavioral effects in children is not explained by OEHHA.. "

Response 3.8: The role of purinergic receptors in brain and their involvement in neurological disorders is a recent and rapidly advancing area of investigation. Just recently (2019/2020), associations with Alzheimer's disease, developmental neurogenesis, multiple sclerosis, substance abuse, depression and schizophrenia have been uncovered. A role for purinergic receptors in children's neurobehavioral disorders has not as yet been studied but cannot be excluded. As noted above, OEHHA conducted a thorough review of the literature seeking information on potentially relevant neurological impacts, and did not limit studies based on what is currently known about pathways of neurobehavioral effects.

Comment 3.9: (page 2) studies on protein binding excluded by Chappell et al. (2020): "(Chen et al., 2016) was not identified in the search conducted by Chappell et al. (2020) because studies on protein binding without measures of neuro-relevant mechanistic events were not included in the objective of that analysis."

Response 3.9: Chen et al. (2016) studied protein binding in connection with amyloid aggregation, which has been identified as a potential mechanism of Alzheimer's disease. This is clearly "neuro-relevant" although apparently not by Chappell et al. criteria.

3.2.4. Several comments below relate to the general IACM comment that OEHHA expanded literature review for dyes with scant or no neurotoxicity data

Comment 3.10: (page 4) "Seven studies identified by OEHHA (Section 4.1.6, page 144) provide toxicology information from oral exposure routes (Hollingsworth, 1982; Kobylewski and Jacobson, 2012; Borzelleca and Hogan, 1985; Borzelleca et al., 1985; Butterworth et al., 1975; Gaunt et al., 1969; Hansen et al., 1966a). These studies were either not identified within searches conducted by others[2/3/] – for example, Hollingsworth, 1982, not peer-revied (sic) or publicly available; Kobylewski and Jacobson, 2012, review article cited in OEHHA for treatment effect on mammary tumor incidence; Butterworth et al., 1975, general toxicity, cited but not detailed in OEHHA report – or were identified but excluded from the assessment because they did not provide neuro-related mechanistic data (i.e., Borzelleca and Hogan, 1985; Borzelleca et al., 1985; Gaunt et al., 1969; Hansen et al., 1966a – all short term or chronic toxicity/carcinogenicity studies without mention of neuro-related effects).

"While OEHHA considers an "indication of possible Blue No. 2 neurotoxic effects is the production of brain tumors" (Section 4.1.6, page 144) based on the significant increase in gliomas in rats fed Blue No. 2 in the diet (2% of diet), without relevant discussion or data linking mechanisms of brain tumorigenicity to mechanisms of neurobehavioral changes, the relevance of the tumor data to neurobehavioral effects in children is unclear"

Response 3.10: Due to lack of neurotoxicity studies in the published literature, OEHHA dug deeper for information. No neurotoxicity studies were conducted in connection with dye registration, but we tried to examine the "lifetime toxicity/carcinogenicity" studies for any indications of neurotoxicity. FDA did not provide OEHHA with copies of these studies and they are only briefly described in peer-reviewed publications (Borzelleca and Hogan 1985; Borzelleca et al. 1985; Butterworth et al. 1975; Gaunt et al. 1969; Hansen et al. 1966a). Hollingsworth et al. (1982) was obtained by FOIA request from FDA and Kobylewski and Jacobson (2012) contains other information obtained by FOIA. Identification of brain tumors produced by Blue No. 2 was a finding considered worth reporting from this effort.

Comment 3.11: (page 4) Blue No. 2. "Five studies were identified by OEHHA (Section 4.1.6, pages 144–145) based on the use of Blue No. 2 for color-based visualization in clinical diagnostics, colonoscopy, and brain tumor surgery (Erickson and Lauron, 1960; Jo et al., 2013; Lee et al., 2015; Choi et al., 2011; Kawaguchi et al., 2007). These studies administered Blue No. 2 through non-oral routes; not considered in other's assessments because the inclusion criterion for the route of administration was limited to oral exposures only.[3]"

Response 3.11: In searching deeper into the literature for biological actions of Blue No. 2, we read studies on clinical applications of this dye. Cardiac side effects potentially mediated by serotonin and histamine could link to brain effects mediated by these neurotransmitters. The "non-oral" routes of administration produced systemic exposures that would have to be referred back to oral routes if detailed pharmacokinetic studies in humans were available.

Comment 3.12: (page 6) Ashour and Abdelaziz et al (2009). "OEHHA identified one single short-term toxicology study in rats (125 mg/kg/day by gavage) (Ashour and Abdelaziz, 2009) (Section 4.1.5, page 142); however, this study did not provide mechanistic data that mapped to neurobehavioral adverse outcomes and therefore, was not included in other assessments of the evidence.[3/]

Response 3.12: The meager pharmacokinetic, toxicology, and mechanism research available for Green No. 3 led OEHHA reviewers to look in detail at this recent study that was conducted

at a dose below the NOAELs used for ADI development. The study suggests that Green No. 3 is biologically active when administered orally at this low dose, and that further studies of neurotoxicity would be valuable. Restricting the review to papers that map to "neurobehavioral adverse outcomes" is extremely limiting, as most of the neurobehavioral adverse outcome pathways are not known.

Comment 3.13: (page 5) "Two *in vitro* studies (Shinoda et al., 1999, inhibition of human aldehyde reductase; and Kuno and Mizutani, 2005, study of a drug-metabolizing cytochrome P450) from the literature were discussed in the OEHHA assessment (Section 4.1.6, page 145) but were not identified by Chappell et al. (2020), because they were outside the scope of the assessment. These studies did not report data considered specific to neuro-relevant pathways. The rationale for including these studies and/or a relationship between inhibition in human aldehyde reductase and neurobehavioral effects is not given in the OEHHA report".

Response 3.13: These studies were included in a paragraph on the metabolism of Blue No. 2. Chapter 4 of the OEHHA review includes toxicokinetics and mechanism. Inclusion of these studies is relevant to toxicokinetics.

Comment 3.14: (page 6) Van Hooft (2002): "Two *in vitro* studies from the literature (van Hooft, 2002, study on hippocampal synaptic function; Yang et al., 2019, study on purinergic receptors) were discussed in the OEHHA assessment (Section 4.1.5, page 143), but were not included in Chappell et al. (2020).[2]/ The study on synaptic function used Green No. 3 as a dye to identify proteins in neuronal cell cultures."

Response 3.14: This study was conducted to determine whether Green No. 3 was neurotoxic when used as a dye in neuronal cell culture. Effects were found on frequency of conduction. This reflects the diligence of *in vitro* investigators, as well as clinical investigators, in looking for "side effects" of dyes when used to stain proteins. OEHHA captured this source of information in its in-depth review.

Comment 3.15: (page 9) Kurebayahsi et al.: "OEHHA (Section 4.1.1, page 135) referenced two *in vivo* studies on thyroid effects that they considered mechanistic, both of which TS excluded because they either did not specifically map to neuro-relevant pathways or they did not address one of the seven FD&C synthetic colors…The Kurebayashi et al. study included exposure of mice to Rose Bengal, Food Red No. 105, without any mention of erythrosine or Red No. 3"

Response 3.15: This study included description of a similar study from the same lab of Red No. 3 (Minegishi et al. 1986) which we have not been able to obtain in translation. Kurebayashi et al. will now be removed from the final version of the OEHHA review.

Comment 3.16 (page 8) Shimizu et al. 2013: "Shimizu et al. (2013) reported Red No. 3 inhibition of iodotyrosine deiodinase *in vitro*. This study was not discussed in any detail within the OEHHA assessment. …this inhibition of iodotyrosine deiodinase could not be bridged to any adverse outcomes associated with exposure to Red No. 3 based on a series of *in vivo* studies conducted in rats and as reviewed elsewhere (summarized in Table S8-1 in Supplement 8 in Gentry et al., *submitted*).3/)"

Response 3.16: OEHHA is not able to review the "series of in vivo studies" because Gentry et al. (submitted) is not publicly available.

Iodotyrosine deiodinase is a major enzyme in the degradation pathway of thyroglobulin, the precursor of thyroid hormone. It releases iodine for re-use in thyroid hormone synthesis. Lower activity of this enzyme leads to hypothyroidism. This is the pathway described by the authors as potentially relevant to Red No. 3 thyroid effects.

Comment 3.17: (page 8) Dalal and Poddar (2009, 2010): "The OEHHA assessment also notes the transient nature of the changes - Section 3.3.5, page 102, it is noted that the changes in neurotransmitter levels were transient, reducing its toxicological status."

Response 3.17: The changes in neurotransmitter levels were transient, which enhances their relevance to the time-limited effects seen in children after challenges with food dyes. Integration of animal and human studies is one of the features of the OEHHA review which is not undertaken by TS. Because an effect is transient is not a reason to "reduce its toxicological status". For example, many chemicals have transient acute effects but are considered toxic.

Comment 3.18: (page 13) Gao et al. (2011), Mohamed et al. (2015): "In referring to the Mohamed et al. (2015) study, it was stated, "As in the Gao et al. (2011) study, generation of ROS [reactive oxygen species] by Yellow No. 5 metabolites was suggested as the mechanism of the effects." However, Gao et al. (2011) was identified as reporting measures of oxidative stress in the brains of rats, including increased ROS, only at the highest bolus gavage dose levels (250 and 500 mg/kg bw/d, or 25-50x ADI).[3]/ Gao et al. (2011) and Mohamed et al. (2015) reported purity of the test substance; however the high dose levels administered by bolus oral gavage may be of concern based on the significance of these finding under a exposure scenario that is of questionable relevance to humans."

Response 3.18: These doses are as relevant to human exposure as the animal doses used to determine the FDA ADI of 5 mg/kg/d. In deriving ADIs, a 100-fold safety factor was used in converting animal NOAELs to ADIs. The ADI based on the 125 mg/kg/d NOAEL in Gao et al. (2011) would be 1.25 mg/kg/d, lower than the current ADI. In addition, a bolus oral gavage is similar to consuming a drink with synthetic food dyes, as is commonly experienced by U.S. children.

Comment 3.19: (page 14) Bhatt et al. (2018): "Bhatt et al. (2018), as referenced in OEHHA Section 4.1.2, page 137, measured markers of oxidative stress in the brain of administered rats. As stated by OEHHA, "The investigators attribute the brain effects to generation of reactive oxygen species (ROS) by Yellow No. 5 aromatic amine metabolites." While it is known that Yellow No. 5 is metabolized in the gut to aromatic amine sulphanilic acid, Bhatt et al. do not provide direct evidence of the presence or activity of sulphanilic acid. Although the study was conducted at the EFSA ADI (7.5 mg/kg/day), this was the only dose level evaluated using a test article in which purity was not provided; also, methodological quality (study reliability) was considered to be low according to SciRAP.[3]/ Without corresponding measures of behavioral activity, the relationship between changes in measures of oxidative stress in the brain and potential adverse neurobehavioral outcomes cannot be established."

Response 3.19: An oral dose at the EFSA ADI was used in this study. Testing at doses identified as safe in humans is important and was the goal of the study. The lack of behavioral

measures is typical of mechanism studies. Of note, Gao et al. (2011) found effects on behavior and oxidative stress with the same LOAELs and NOAELs.

While Bhatt et al (2018) did not measure the metabolites, sulphanilic acid is the main metabolite of Yellow No. 5.

Comment 3.20: (page 14) Rafati et al. (2017) is "mentioned in Section 7.3.1.2, on page 257, where it was indicated that brain histomorphometric endpoints (dendritic spine length and brain effects) were used to identify the LOAEL of 5 mg/kg/day.,,, Thus, interpretation of findings was complicated by significant study limitations, including lack of reporting of test substance purity information, use of rats with advanced age, and administration via oral gavage versus dietary exposure".

"OEHHA (Section 5.3.4, page 166) noted that "The brain assays demonstrated alterations in cell number, volume, and cell shape in the medial frontal cortex in dye-treated animals compared to controls. However, it was not clear what "dye" the authors were referring to".

Response 3.20: This study of Yellow No. 5 is rich in rare data on brain histomorphometry effects of food dyes. As noted above, the rats were not of advanced age (this was a misreading of the study by Ramboll) and gavage administration is a relevant parallel to the challenge studies that identified food dye effects on behavior in children.

The OEHHA text clearly identifies Yellow No. 5 as the dye referred to:

"In one study of Yellow No. 5 (Rafati et al. 2017) more days were required for Yellow No.5 treated rats (low- and high-dose groups combined) to reach the learning criterion in the radial arm maze. More errors during learning and during the retention phase were observed in the dye-treated groups. The brain assays demonstrated alterations in cell number, volume, and cell shape in the medial frontal cortex in dye-treated animals compared to controls."

Comment 3.21: (**Appendix B, page 16, 17; IACM comment letter, page 19, 20)** Osman et al. (2002,2004):

3.21a. "One mechanistic *in vivo* study identified by both TS and OEHHA (Osman et al., 2004) examined cholinesterase (ChE) enzyme activity following dietary exposure to Yellow No. 6 (or its metabolite sulphanilic acid) at only one dietary exposure concentration estimated to be ~4 g/kg bw/d (which is the ADI for the parent compound) (detailed in Table S8-4 in Supplement 8 in Gentry et al. (submitted)"

"There were statistically significant changes in ChE activity following exposure to both the parent compound and its metabolite. OEHHA noted (Section 4.1.4, page 142) that the data suggested a higher potency for inhibiting ChE for sulphanilic acid compared to the parent compound. Potency is difficult to evaluate *in vivo*, especially with data collected at only one dose level."

"The OEHHA assessment (Section 4.1.4, page 142) stated, "This finding with sulfanilic acid suggests it is the active agent for Yellow No. 6 effects on cholinergic systems, as well as for effects on behavior (Goldenring et al., 1982) and identif[ies] this neurotransmitter system as a potential mechanistic pathway for Yellow No. 6 neurotoxicity." However, Osman et al. (2004) is of low quality/reliability based on the SciRAP assessment and the lack of information on test-article purity, along with being conducted at only one exposure concentration,[3]/ thereby

precluding any dose-response evaluation"

"Although the *in vivo* study described by Osman et al. (2004) was of poor quality, it supports previous *in vitro* findings reported by the same investigators (Osman et al., 2002) in which Yellow No. 6 inhibited human cholinesterase and pseudocholinesterase activity, discussed by OEHHA (Section 4.1.4, page 142).

Response 3.21a: These two research reports document a series of investigator-initiated, hypothesis-testing studies of Yellow No. 6 inhibition of cholinesterase.

In Osman et al. (2002) the investigators first identified Yellow No. 6 cholinesterase (ChE) inhibition *in vitro* with dose-response studies in human blood samples, and distinguished between effects on red blood cell ChE and plasma ChE. This is an important distinction because plasma ChE is a validated marker of human poisoning by insecticides that inhibit ChE. They also determined that the cholinesterase inhibition was reversible.

In Osman et al. (2004) ChE inhibition by Yellow No. 6 was compared to its two main metabolites, sulfanilic acid and naphthionic acid, using the same *in vitro* assay. This is important because it is known that Yellow No. 6 is extensively metabolized in the gut, with metabolites absorbed into the systemic circulation in much larger amounts than the parent compound. Finally, moving from *in vitro* to *in vivo* studies, they determined that this effect occurred with dietary dye administration. Rats and were fed Yellow No. 6 (or its metabolites) in the diet at a dose equivalent to the current JECFA ADI.

TS based some of its comments on its guidelines and criteria for including or excluding studies for use in quantitative risk assessment, citing Gentry et al, (submitted). OEHHA is not able to review the SciRAP evaluation because Gentry et al. (submitted) is not publicly available.

Osman et al. refer to a previous publication from their lab for details of the methods. These need to be retrieved for a full evaluation of study quality. Journals often limit the size of articles and suggest citing previous work for details. The single study at the ADI dose was adequate for hypothesis-testing and provides valuable information about cholinesterase inhibition without a full dose-response study.

OEHHA assumes TS meant 4 **mg**/kg bw/d (this is the current JECFA ADI) in referring to the dose level in Osman et al (2004).

3.21b. "Although TS did not review or conduct a quality assessment on the *in vitro* study (Osman et al., 2004), it was noted that there was no positive control used in this assay, along with a lack of confirmation of test substance solubility and/or toxicity in the model system evaluated (Osman et al., 2002)."

"OEHHA (Section 5.3.5, page 168) goes on to state, "Studies with sulfanilic acid, a metabolite of Yellow No. 6, [suggest that] it is the active agent for Yellow No. 6 effects on cholinergic systems, as well as for effects on behavior[,] and identify this neurotransmitter system as a potential mechanistic pathway for Yellow No. 6 neurotoxicity." Both of these OEHHA conclusions are based on one poor-quality *in vivo* study (Osman et al., 2004), and an *in vitro* study (Osman et al., 2002, discussed above)." (IACM comment letter, page 19,20),

Response 3.21b: Osman et al. (2004) contains an *in vitro*, as well as an *in vivo,* study. This study confirmed the Yellow No. 6 effect on cholinesterase, the reversibility of the effect, and the relative potency of the parent compound and the metabolites. We disagree that the *in vivo*

study was of poor quality. It achieved its goal in this hypothesis-testing series of experiments of demonstrating that the effect previously studied *in vitro* could occur *in vivo*. While Osman et al. is described as "an *in vitro* study" it is actually a series of *in vitro* studies using human plasma and red blood cells and demonstrating dose-response, computing Lineweaver-Burke plots and dissociation constants, identifying true and pseudo-cholinesterase activity, and confirming reversibility through use of dialysis. The Ellman et al. cholinesterase assay used in these papers was recently described as "the most used and adopted method for cholinesterase activity identification" in a 2012 review of progress in cholinesterase quantification methods. In sum, these two papers represent a valuable, high quality examination of a potential pathway for Yellow No. 6 neurotoxicity.

The TS analysis of the Osman et al. studies points out the inadequacy of the tool-based screening of studies for inclusion/exclusion in quantitative risk assessment as a useful approach to scientific review. A comprehensive review to assess potential health hazards needs to be as thorough as possible and not focus just on dose-response assessment as TS has done. There are many areas of research that inform the hazard identification that TS missed because of their use of SciRAP and focus on dose-response. It is not surprising that they came to a different conclusion based on an overly constrained review. As stated by Klimisch et al.(1997):

"*All available experimental data as compiled in a data sheet (IUCLID) should be considered in risk assessment because only the totality of data will increase clarity of the conclusions. Thus limitations of publishing only a ''definitive data set'' appear to limit clarity and worldwide understanding. The relevance and adequacy of all the data used in a risk assessment process should be defined by expert judgment in a comprehensive report.''*

Comment 3.22: (Appendix B, pages 11-12) Noorafshan et al. 2018: "The purity was not reported in this study; as such, it is possible that the effect at the high dose could be attributed to impurities in the test article."

"OEHHA states that, in this study (Noorafshan et al., 2018), "sample size was small, particularly for the histomorphology/stereology (6 animals per dose group)" (OEHHA, Section 7.5.2, page 273)." This would suggest a concern for confidence in the study findings."

"OEHHA states ""However, there is mechanistic support for oxidative damage from Red No. 40 from other studies and the anti-oxidant taurine reportedly reversed the effects of Red No. 40. Additionally, the changes in the medial prefrontal cortex can be directly related to the cognitive performance of the animals, as this part of the rodent brain is involved in spatial memory, decision-making and attention (Noorafshan et al, 2018)." The specific "other studies" that OEHHA was referring to were not referenced."

Response 3.22: Sample purity is not addressed in this study. Impurities could always be responsible for test results at any dose unless the sample is 100% pure.

The sample size is well within range for histomorphometric studies. A smaller sample size increases the chance of false negatives but does not reduce confidence in study findings.

The reference for Red No. 40 oxidative damage effects is Khayyat et al. 2018. It will be included in the final document.

Noorafshan et al. (2018) was the only Red No. 40 study reviewed by TS in Appendix B although it was not identified as relevant by Chappell et al. (2020). Four other Red No. 40

studies were reviewed by OEHHA but excluded by TS as "out-of-scope" or "evaluation of a mixture of colors". As noted in our responses above, dismissing animal or human studies that assessed the potential mechanisms of mixtures of food dyes throws out information related to whether the synthetic food dyes affect behavior. As well, what was "out of scope" to TS was not out of scope for OEHHA's purpose of evaluating all available scientific information relevant to whether food dyes potentially impact neurobehavior in children.

4 Response to Comments on the OEHHA ToxCast Analysis

In the following sections, OEHHA summarizes the substantive issues raised in the IACM comment letter, including Appendix B, written by ToxStrategies, regarding the ToxCast HTS data and provides responses to these issues. Several issues are repeated throughout the IACM summary and Appendix B, therefore, OEHHA consolidated many of the comments.

In IACM's comments pages 17-20, they conclude that the HTS data do not support an association between exposure to any of the dyes and potential neurobehavioral effects. For example, the text on page 17 of the IACM comments reads "Taken together, there was insufficient evidence to support an association between exposure to Blue No. 1 and mechanistic changes that are potentially related to neurobehavioral outcomes." This or similar comments are included for each of the seven food dyes. OEHHA disagrees with these conclusions, which is based on a more limited analysis of the HTS assays, published in Chappell et al. 2020, than that conducted by OEHHA. We discuss the reasons for our disagreement in the responses to comments below.

Comment 4.1: (IACM comment letter, page 14) "While the OEHHA report mentions various *in vivo* and *in vitro* studies of mechanistic data for each color, there is no overall conclusion regarding how such data contribute to the understanding of underlying mechanisms of neurobehavioral effects, nor are such data considered in the context of plausible modes of action or adverse outcome pathways related to neurobehavioral outcomes. There is an overall lack of commentary on dose (relevance to human exposure), duration/timepoints, models used (predictability or consistency of models used, either *in vivo* or *in vitro*), or "bridging" of molecular or cellular signals to behavioral effects in the OEHHA report. It is important to note that a single signaling event is not enough for an adverse outcome pathway to be relevant, but rather multiple key events (KEs) are required to get from the molecular initiating event (MIE) to the adverse outcome."

"The OEHHA report considered the high-throughput screening (HTS) assay data to provide 'limited support for *in vivo* 'neurotoxicity,' based on the fact that the colors had some, albeit inconsistent, activity in HTS assays deemed relevant to neurobehavioral effects. Further, clarity was not provided on how such *in vitro* data were specifically related to outcomes in humans and/or laboratory animals. Although the OEHHA report included an extensive assessment of HTS data, such as bioactivity of HTS assays, from the ToxCast/Tox21 program for the seven colors, the analyses included many non-specific assays for neurobehavioral effects (e.g., markers of oxidative stress or inflammatory response in non-neuronal cell lines), and did not account for confounding by refraining from integrating potential cytotoxic interference or other data quality issues (such information is provided in the ToxCast/Tox21 database). OEHHA did not attempt to elucidate underlying mechanisms for neurobehavioral effects with the use of HTS data. Instead, the report concludes generally that the HTS data are limited yet supportive of *in vivo* neurotoxicity observations for the food colors, without specifying connections between HTS results and *in vivo* outcomes."

Response 4.1: OEHHA integrated evidence from multiple data streams in considering the potential neurological activity associated with the food dyes, including information from animal bioassays, data from genotoxicity and metabolism studies, and data from epidemiological studies. A small section of this comprehensive review reported on HTS data. The intent of the exploration of the HTS data, was to assess potential underlying molecular mechanisms that may be linked to neurological pathway perturbations. The limited capability to extrapolate the *in vitro* hit data to *in vivo* observations of specific and nonspecific toxicity is not unexpected.

OEHHA states in the report that while there are no clear mechanisms by which any of the dyes induce neurobehavioral effects, there are some insights provided by both *in vivo* and *in vitro* data supporting potential mechanisms. We discuss these potential pathways in Chapter 4. The difficulty in using HTS to extrapolate from *in vitro* bioactivity to *in* vivo effects is well known, with adverse outcome pathway (AOP) frameworks often proposed as a means to help provide this linkage. However, the idea that the mechanism can only be substantiated by the development of an AOP is so constraining as to be impractical - there are generally very little data for most chemicals to clearly delineate an adverse outcome pathway, and there may be several pathways that lead to the same measured outcome and are not yet known. OEHHA will revise Section 4.3.4 to integrate more information on the interpretation of AOP in this current framework as well as the limitations of their application. An AOP is not required to associate a chemical with a downstream effect.

Contrary to what is implied in the comment, in Chapter 4 and Appendix A, OEHHA describes in detail the method we used to determine which HTS assays to evaluate in our report and why we chose those assays as potentially relevant to neurological endpoints. As noted in the OEHHA Report (section 4.3.2): “Initially, OEHHA screened the food dyes in publicly available aggregate databases including the Comparative Toxicogenomics Database (CTD; (Davis et al. 2019)), Chemical Hazard Data Commons (CHDC), and the Chemistry Dashboard (Williams et al. 2017) to evaluate whether there were any known established associations between the food dye chemicals and neurological process targets linked to toxicity *in vivo*. In particular, OEHHA evaluated the food dyes in the Chemistry Dashboard in seven developmental neurotoxicity (DNT) lists to see if there were any hazards already associated with the food dyes. Due to a lack of available information, OEHHA developed an approach to map potential associations between the food dyes and neurological activity based on existing *in vitro* data. This approach incorporates a strategy for linking the potential molecular targets examined in assays to neurological processes and used chemicals with known DNT endpoints to profile potential neurological markers. There were 108 NVS assays, 50 ATG assays, and 24 Tox21 assays selected based on whether the assays: 1) had a neurological-related gene target; 2) were conducted in brain tissue (regardless of species); or 3) targeted the specific receptors of aryl hydrocarbon, androgen, estrogen, or the thyroid hormone, based on literature reporting interactions between food dyes and these receptors (Axon et al. 2012; Dees et al. 1997; Jennings et al. 1990; Mathieu-Denoncourt et al. 2014). Such interactions may have downstream effects on targets underlying neurological processes, and therefore, these assays are pertinent to explore as well. There were 182 total assays from these three criteria.”
Although non-neuronal cell lines were investigated in this study, one purpose of our work was to use HTS assays to explore whether there was concordance with what has been established in

literature. There are two papers on the oxidative stress in the brain after *in vivo* treatment of Yellow No. 5 of rats (Bhatt et al. 2018; Mohamed et al. 2015). Oxidative stress is viewed as an imbalance between the production of reactive oxygen species (ROS) and their elimination by protective mechanisms, which can lead to chronic inflammation and thus cellular and tissue distress – therefore it was pertinent that we utilized assay sets mapped to both these mechanisms. The OEHHA report addresses the scope of our work as well as the limitations of the HTS assays in covering endpoints related to the key characteristics of neurotoxicants. The limited support provided by the HTS data is discussed on page 160.
The integration of potential cytotoxic interference or other data quality issues are addressed in OEHHA's response to comment 4.2.

Comment 4.2: (Appendix B, pages 21-22) Consideration of Cytotoxicity and data quality flags for HTS assay activity. "OEHHA (where stated) applied no filter for cytotoxic interference (i.e., the AC50 value of activity relative to the cytotoxic concentration), nor data quality flags (assigned by ToxCast, related to issues in data analysis and model fitting), nor chemical analytical quality control (Q.C.) (e.g., purity and identity). Determination of activity or inactivity was based solely on the "hit-call" provided in the ToxCast database. While Judson et al. (2016) is cited by OEHHA as stating that the cytotoxic burst should not be used as a filter, cytotoxic interference is a known and well- established factor that should be considered in the interpretation of *in vitro* data as discussed in the very same paper (Judson et al., 2016)."
"Outside of cytotoxic burst criteria, the viability assays in the Tox21 program are specific to measuring cell death related to specific individual assays. There is no reason to ignore such assay information that would otherwise help with appropriate analyses and interpretation of the results."

"OEHHA did not consider data quality flags in their assessment (page 146, Section 4.3.1). Chappell et al. (2020) noted more than one data quality flag would render the assay not "active".
"OEHHA makes no mention of chemical quality information in Section 4.3 on HTS assays. While Chappell et al. (2020) did not exclude data based on sub-optimal or absent chemical quality data, such information was discussed in the broader context of overall interpretation of the data."

"It should be noted that other groups/programs are applying various criteria when deciding when to assign a classification of "active" to responses observed for HTS assay endpoints, and/or to include the assay data in an assessment.

- Relative to National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM)/Integrated Chemical Environment (ICE) curated high-throughput screening (cHTS) data:
  - Data quality is considered by omitting any assay endpoints with a "hit-call" of active from curated data when: the assay is a down-direction assay (i.e., inhibition, antagonism, loss-of-signal, etc.), and the best-fit curve was a gain-loss model; or the best-fit curve was a gain-loss model, and only a single mid- range concentration had activity above the activity cutoff threshold, among other criteria.
  - Sample quality is considered by omitting assay endpoints from the curated ToxCast HTS data within ICE in which chemicals with a chemical QC grade of "caution" were used, among other criteria.

- "Interestingly, and somewhat contradictory to the approach taken for synthetic colors, OEHHA itself excluded HTS assays based on data quality flags and/or chemical Q.C. issues (i.e., lacking Q.C. information or major issues with chemical analytical Q.C. results) in their evidence on the Carcinogenicity of Acetaminophen (2019). "

Response 4.2: We agree that data flags increase uncertainty and complicate interpretation. We disagree that one should simply dismiss those results and we incorporated all the results into our analysis regardless of data flags. Judson et al (2016) states "However, given quantitative uncertainties in AC50 values, and our incomplete understanding of the biological linkages between cell-stress, cytotoxicity, and specific activity, a hard filter will not always be appropriate." and in personal communication with the authors of that study ((R. Judson, personal communication, April 3, 2020), flags and cytotoxicity limits should not be used automatically as hard filters, but should be considered when evaluating and interpreting HTS results regarding the degree of uncertainty that may be associated with the HTS results. In addition, flags are subject to change as the ToxCast data analysis pipeline evolves. Our results were extracted from the iCSS Dashboard in 2019. Since then, much of the data has been re-pipelined and new data has also been incorporated. This could have resulted in changes to the parameters that contribute to defining the hit call, which most often affects borderline hits and flags associated with them. Thus, it stands to reason that flags may change in different and future versions of ToxCast.

Further, there is not an established standardized way of incorporating cytotoxicity when interpreting ToxCast HTS data. Viability assays were given limited consideration in our evaluation due to a number of factors including 1) the variable number of cytotoxicity assays for each chemical; 2) appropriateness of utilizing Tox21 cytotoxicity assays in application to ToxCast assays outside the Tox21 platform (potentially important differences in the methods across platforms which would affect cell viability and non-concurrent measurements); 3) many cytotoxicity assays have flags of efficacy, which by the same rules as used by ToxStrategies, would render them less reliable. At the very best, an active hit at a concentration below the cytotoxicity limit indicates that the active hit is <u>not likely</u> due to processes involved in cell death (R. Judson, personal communication, April 3, 2020). However, the inverse, that active hits above the cytotoxicity limit are likely due to processes involved in cell death, is not necessarily accurate and cannot be applied to justify ignoring those assays. It may cast uncertainty over the result, but we do not know in fact that the result is due to processes involved in cell death. Further language regarding rationale for not using cytotoxicity as a hard cutoff has been incorporated in Chapter 4 and Appendix A.

Given these issues, OEHHA will continue to follow the recommendation by Judson et al. in that rather than using data flags as hard filters for the relevancy of assay results, users can benefit from understanding that the flags for each assay may identify potential interference and possible sources of uncertainty. It is clear flags should not be used to determine whether an assay is active. Using flags and cytotoxicity as hard cutoffs to ignore assay results would exclude potentially important information.

Appendix D in OEHHA's report summarizes the HTS assay results for the different dyes and has a column where the flags are all noted for each dye. We chose not to ignore the results of the HTS assays even where there were flags after the recommendation in Judson et al (2016),

and after discussion with the staff at U.S.EPA. We also did not make independent "hit-calls" as Chappell et al., (2020) did, filtering out results with data quality flags as inactive.
There is not yet a standardized or tiered approach for integrating cytotoxicity and flags of efficacy when assessing the ToxCast HTS data.

A section on the chemical quality of the different food dyes is now included in the revised draft report- Appendix A. In terms of chemical analytical quality control, Blue No. 1, Red No. 3, and Yellow No. 5 have caution flags regarding their quality control. The results of assays on these dyes should not be excluded solely on caution flags; instead, the assay results should be assessed with a consideration of uncertainty. Although there may be numerous sample IDs for the same food dye color, the Chemistry dashboard only provided chemical QC data on food dyes that had a Tox21 ID. The following information on chemical quality is available from the Chemistry Dashboard:

Dye	Name	CASRN	DSSTox Substance ID	Tox21 ID	QC Grade
Blue No. 1	Brilliant Blue	3844-45-9	DTXSID2020189	Tox21_300516	T0; (Z) MW Confirmed, no purity info T4; (Z) MW Confirmed, no purity info
Blue No. 2	Indigo Carmine	860-22-0	DTXSID1020190	Tox21_113456	T0; (ND) Not Determined, Analytical analysis is in progress T4; (A) MW Confirmed, Purity > 90%
Blue No. 2	Indigo Carmine	860-22-0	DTXSID1020190	Tox21_302732	T0; (C) MW Confirmed, Purity 50-75% T4; (Fns) CAUTION, No Sample Detected; Biological Activity Unreliable
Green No. 3	Fast Green	2353-45-9	DTXSID3020673	Tox21_302086	T0; (A) MW Confirmed, Purity > 90% T4; (A) MW Confirmed, Purity > 90%
Red No. 3	Erythrosine	16423-68-0	DTXSID7021233	Tox21_202932	T0; (ND) Not Determined, Analytical analysis is in progress T4; (Cc) CAUTION, Low Concentration, Concentration 5-30% of expected value
Red No. 3	Erythrosine	16423-68-0	DTXSID7021233	Tox21_302085	T0; (ND) Not Determined, Analytical analysis is in progress T4; (F) CAUTION, Incorrect MW Biological Activity Unreliable

Dye	Name	CASRN	DSSTox Substance ID	Tox21 ID	QC Grade
Red No. 40	Allura Red	25956-17-6	DTXSID4024436	Tox21_300393	T0; (A) MW Confirmed, Purity > 90% T4; (A) MW Confirmed, Purity > 90%
Yellow No. 5	Tartrazine	1934-21-0	DTXSID1021455	Tox21_113411	T0; (F) CAUTION, Incorrect MW Biological Activity Unreliable T4; (I) ISOMERS, two or more isomers detected
Yellow No. 5	Tartrazine	1934-21-0	DTXSID1021455	Tox21_201539	T0; (A) MW Confirmed, Purity > 90% T4; (F) CAUTION, Incorrect MW Biological Activity Unreliable
Yellow No. 5	Tartrazine	1934-21-0	DTXSID1021455	Tox21_300554	T0; (A) MW Confirmed, Purity > 90% T4; (Ac) CAUTION, Low Concentration Concentration 5-30% of expected value
Yellow No. 6	Sunset Yellow	2783-94-0	DTXSID6021456	Tox21_201897	T0; (A) MW Confirmed, Purity > 90% T4; (A) MW Confirmed, Purity > 90%
Yellow No. 6	Sunset Yellow	2783-94-0	DTXSID6021456	Tox21_300407	T0; (A) MW Confirmed, Purity > 90% T4; (A) MW Confirmed, Purity > 90%

Lastly, relative to NICEATM/ ICE curated high-throughput screening (cHTS) data, the provided curation standard has not been widely adopted; these curated standards were not used in Chappell et al. (2020). Currently, there is no established method for discarding data based on these criteria. OEHHA will note these curation considerations in future assessments of HTS datasets, however eliminating active hit calls based on fit models or caution QC flags may lead to excluding potentially relevant data. Further supporting evidence should be given to assess these curation criteria.

In reference to OEHHA's 2019 draft document on Acetaminophen, there is not an established standardized way of incorporating cytotoxicity or flags when interpreting ToxCast HTS data. As stated before, the ToxCast data pipeline is continuously evolving and changing, as new data is being incorporated. This could have resulted in changes to the parameters that contribute to defining hit calls.

Comment 4.3: (Appendix B, page 19) "OEHHA Section 5.4, page 174: "Based on the subset of assays we evaluated here, the ToxCast assay results provide limited support for *in vivo* neurotoxicity observations for the food dyes. It should be noted that the assays explored here are intended to provide initial information about the capacity to associate *in vitro* work with the ability for a food dye to promote a biological response. However, these assays are limited for predicting long term or indirect adverse effects in complex biological systems, in part, due to the complexity of the *in vivo* pathway interactions leading to neurotoxicity (including neurobehavioral effects) and DNT [developmental neurotoxicity] compared to the current limited spectrum and range of the ToxCast assays. Evaluation of these chemicals in future iterations may offer more refined results and validate that these gene markers play a critical role in chemically-induced mechanisms of neurotoxicity."

"While we agree with the notion of more HTS testing, any response noted in HTS assays would ultimately only help to formulate hypotheses that would require further testing to determine whether support for a particular adverse neurobehavioral outcome pathway exists. HTS assays represent a good screening and prioritization tool but cannot be used to 'validate' any particular marker as critical to chemically induced mechanisms of neurotoxicity. Additionally, at this time, these individual assays can only be reliably used when information on assay interference such as cytotoxicity and chemical purity and stability is available and incorporated into the assessment… Chappell et al. (2020) concluded that "the results of our assessment of available *in vitro* mechanistic data collected from assays that measure signals related to MIEs or KEs involved in neurodevelopmental processes indicate that the seven FDA-approved food colors (when batch certified) have limited or no activity for such signals. While available information on FD&C colors and genes or enzymes that may have a role in mechanisms of neurodevelopmental alterations may be limited, FD&C Red No. 3 was the only color (of the seven assessed) that showed activity associated with neurodevelopmental pathways."

Response 4.3: OEHHA agrees that the HTS results are in some respects hypothesis generating. However, the results can be used to provide insight into underlying mechanisms of observations in in vivo animal toxicology studies and epidemiology studies. OEHHA notes there are several pertinent associations between the dyes and certain molecular targets of interest. As we state in our report (page 160), OEHHA's intent was "to provide initial information on whether the in vitro HTS [high throughput screening] assays could be linked with the ability of the FD&C synthetic food dyes to promote a biological response in the nervous system." However, we caution that these assays "are limited in predicting long term or indirect adverse effects in biological systems, in part due to the complexity of the mechanistic processes that underlie detrimental neurotoxic or neurobehavioral outcomes compared to the current limited spectrum of the ToxCast assays."

The ToxCast assay platforms and options for analyzing the resulting concentration-response data are continually updated and refined, as is the information provided in the CTD, and the ToxPi software functionality. These updates and evolution present moving targets.

OEHHA concurs that bioactivity in ToxCast assays assessing non-apical endpoints is not always echoed by observations of corresponding apical effects in *in vivo* studies. Therefore, we

do not state that the HTS assays "validate" *in vivo* observations of chemicals, but rather provide supporting evidence.

Comment 4.4: (IACM comment letter, page 13; Appendix B, page 19) "Both OEHHA and Chappell et al. (2020) agree on overall limitations associated with the current set of HTS assays in providing a reliably predictive signal for neuro-relevant biological processes generally, and for the colors under review specifically. The overall activity profile for the various colors differs between the TS and OEHHA assessments however for two main reasons:

- Differences in the set of assays included in the assessments, some that do not appear to be directly relevant to the neuro-endpoint in the case of OEHHA.
- Consideration of data quality issues and/or cytotoxic interference, or lack thereof in the case of OEHHA.

"In the OEHHA report (page 158, Section 4.3.3.5), the following statement is incorrect, "Even with the limitations of the *in vitro* data, in contrast to a recent study published by Chappell et al. (2020), our approach resulted in significantly more active assay hits (283 compared to 116 assays)." In fact, a total of 99 HTS assays were mapped to potential mechanisms of neurobehavioral outcomes in Chappell et al. (2020). Across these 99 assays, the "coverage" of the seven colors tested culminated in a total of only 116 assay endpoints….
The number of active "hit-calls" are, consequently, much fewer in Chappell et al. (2020) compared to OEHHA's assessment due to a lower number of overall assays deemed relevant to neurobehavioral outcomes, as well as integration of data quality issues and cytotoxic interferences, which OEHHA did not account for nor integrate into their assessment. Therefore, OEHHA could state the following to ensure a more appropriate characterization of the difference: "Even with the limitations of the *in vitro* data, in contrast to a recent study published by Chappell et al. (2020), our approach resulted in more assay endpoints included in the assessment (283 for OEHHA compared to 99 for Chappell et al. (2020)) and more corresponding active "hit-calls" (350 for OEHHA and 8 for Chappell et al. (2020). These differences could be explained by the fact that (i) we cast the net much wider to include indirect effects that have questionable associations with neurobehavioral outcomes and (ii) we (OEHHA) did not account for data quality issues or assay interference due to cytotoxicity in determining activity calls in contrast to the approach Chappell et al. (2020) took."

Response 4.4: OEHHA concurs that the key differences in approach between our analysis and that published by Chappell et al. (2020) are the influence of data flags and results for cell viability assays on including or excluding HTS assay results as well as OEHHA casting a wider net for HTS assays that may have relevance for an effect on the nervous system.

OEHHA thanks IACM for their suggested revision. The text of the report has been revised to the following: "Even with the limitations of the *in vitro* data, in contrast to a recent study published by Chappell et al. (2020), our approach resulted in more assays being included in the assessment (283 for OEHHA compared to 99 for Chappell et al. (2020)) and more corresponding active "hit-calls" for OEHHA in comparison to Chappell et al. (2020). These differences could be explained by the fact that (i) we cast a much wider net to include indirect effects that may have potential associations with neurobehavioral outcomes and (ii) we did not integrate cytotoxicity and flags of efficacy as hard cutoffs when determining activity calls in contrast to the approach Chappell et al. (2020) took."

It should be noted that Chappell et al (2020) "integrated" cell viability, cytotoxicity burst information, and other data flags by dismissing the results of the HTS assays that had more than one data flag, or where the AC50 value for the molecular receptor was above the AC50 for the loss of the viability measure. The flags and cytotoxicity limits should not be hard cut-offs for excluding the results. (See response to IACM comment #2).

Comment 4.5: (Appendix B, page 20) "…OEHHA limited their identification of neuro-relevant assays to the NovaScreen (NVS), Attagene (ATG), and Tox21 vendors, whereas Chappell et al. (2020) considered assays from all vendors. OEHHA included other vendors in their selection of assays based on activity of known neurotoxicants (pesticides) and assays related to oxidative stress and inflammation according to Iyer et al. (2019)."

Response 4.5: The rationale for our assay selection is stated in the OEHHA report (Appendix A): "ToxCast assays from the NovaScreen (NVS), Attagene (ATG), and Tox21 platforms were selected to assess target binding as an indicator of protein activity, translated as an association between receptor binding and potential effect. We initially explored just these three platforms to demonstrate a proof of concept while maintaining manageability."

The ToxCast assay platforms and options for analyzing the concentration-response data are continually updated and refined, as is the information provided in the CTD, and the ToxPi software functionality. These updates and refinements present a variety of moving targets, so the results and conclusions of even a comprehensive approach to all the platforms may change overtime. Our intent here is to explore an approach that could have utility in understanding potential mechanisms of neurological or neurobehavioral effects of the food dyes, as well as be applied with future improved iterations of ToxCast and other alternative test method data, for improved chemical characterization outcomes. While adding assay platforms to our model would slightly improve coverage, other more important biological coverage gaps would persist, and the outcomes would still be hampered by the limited metabolic capacity of the assays and the lack of coverage of many biological pathways that could lead to an adverse effect.

Comment 4.6: (Appendix B, page 21): "A number of the Attagene assays that OEHHA used in their list of 283 neuro- relevant assays were not included in Chappell et al. (2020) because they are listed as "not developed or optimized to detect loss of signal" in the CompTox database and are only optimized for gain of signal modeling (e.g., ATG_THRA1_TRANS_DN, ATG_GPCR_DRD1_TRANS_dn, among others)."

Response 4.6: OEHHA's rationale for our assay selection is covered in pages 147-148. Differences in our approach from Chappell et al. (2020) are covered in pages 158-159. The example assays (noted in the IACM comments) are inconsequential to our overall results because they were either not tested or inactive for the food dye set.

Comment 4.7: (Appendix B, page 21) "It is worth noting that the U.S. Environmental Protection Agency (EPA), the National Center for Computational Toxicology [now the Center for Computational Toxicology and Exposure] (NCCT/CCTE) – the EPA division that works with

these assays – is not using HTS assays currently for either hazard or risk assessment. Consequently, it is imperative that the cytotoxic interference information is considered in the determination of assay activity; and, more specifically, it provides the necessary context to assign assay data as "unreliable" vs. active or inactive."

Response 4.7: The comment that the U.S. EPA is not using HTS data for hazard identification risk assessment may be true, but it is a bit misleading. Scientists from NCCT have published many papers utilizing results of HTS assays, including comparing to in vivo activities. No one to our knowledge at U.S.EPA has officially used HTS in quantitative risk assessment, although several papers published in scientific journals, using *in vitro* to *in vivo* extrapolation, have compared estimates of current exposures and resulting blood levels to concentrations in cells that cause perturbations. A number of scientists outside U.S.EPA have also utilized HTS data to explore connections to in vivo toxicity, including scientists at CalEPA (Silva et al., 2015, Pham et al., 2017, Iyer et al., 2019). The goal of OEHHA's exploration of the HTS data for food dyes in our report is to provide a thorough review of whether there is any mechanistic evidence supporting the potential for food dyes to interact with neurological targets as a supplemental line of evidence. We do not base our conclusions solely on the results of the HTS assays, nor do we use it for quantitative risk assessment. As mentioned in response to Comment #2, there is not a standardized way of incorporating cytotoxicity. However, OEHHA agrees that instead of using cytotoxicity or flags as filters to determine an assay as active or inactive, the data should be evaluated within the context of "reliability" in view of the flags.

Comment 4.8: (Appendix B, page 6) "Comment on Green No. 3. "The OEHHA assessment, which did not apply filters for either data quality issues or cytotoxic interference in the determination of activity, considered Green No. 3 active in assays across all the neuro-relevant categories identified in their assessment in which the color was tested, as presented in Summary Table 4.1 (pages 152–154)."
"only limited neuro-relevant mechanistic information was identified by OEHHA for Green No. 3. OEHHA states (Section 5.3.5, page 169), "Both Blue No. 1 and Green No. 3 inhibit purinergic receptors"; however, the relevance of this finding to potential neurobehavioral outcomes is unclear because inhibition of purinergic receptors has potential neurotherapeutic value and would be considered not relevant for evaluating adverse neurobehavioral effects."

Response 4.8: As noted earlier, OEHHA did not use flags as hard cutoffs to discard results. Please see OEHHA's response to comment #2.

OEHHA disagrees that any and all actions at purinergic receptors have "neurotherapeutic value". Activity at such receptors is indicative of ability to perturb neuronal tissue, which is what the HTS assays are meant to discover. To ignore a perturbation because it might have positive effects is inappropriate. One can argue that drugs have positive effects, but deleterious effects as well.

Comment 4.9: (Appendix B, page 10) "For ToxCast HTS assays, Chappell et al. (2020) applied filters for assay- specific data quality and cytotoxic interference in the determination of assay endpoint activity. Red No. 3 was active in four of the 15 neuro-relevant assays (as mapped within Chappell et al. 2020). These four assays were active for loss of gene expression

related to the production, transport, or degradation of the neurotransmitters dopamine and serotonin. The OEHHA assessment, which did not apply filters for either data quality issues or cytotoxic interference in the determination of activity and therefore brings into question the applicability of this assessment, considered Red No. 3 active in assays for all neuro-relevant categories in which it was tested (noted in the Summary Table 4.1, pages 152–154). Although the OEHHA assessment specifically highlights activity related to the androgen, estrogen, and thyroid receptors, TPO inhibition, and monoamine oxidase, only the loss of signal for monoamine oxidase by Red No. 3 was reported in Chappell et al. (2020) assessments after data quality considerations were integrated."

Response 4.9: As noted above, OEHHA did not use flags as hard cutoffs to discard results. Please see OEHHA's response to comment #2. OEHHA views activities in the HTS assays related to hormonal receptors, TPO inhibition, and monoamine oxidase as potential pathways for neuronal perturbation. As well, in our report, we summarize several other *in vivo* and *in vitro* studies that measure perturbations in neurotransmitter release and transport and uptake (see section 4.1.1 of OEHHA report).

Comment 4.10: (IACM comment letter, page 17; Appendix B, page 12) "There was no activity in *in vitro* HTS assays reporting neuro-related mechanistic data for Red No. 40. Overall, there was insufficient evidence to support an association between exposure to Red No. 40 and neurobehavioral-related key events."

"In summary, very limited neuro-relevant mechanistic information is available for Red No. 40 to evaluate mechanistic changes as they relate to potential alterations in neuro- relevant pathways that lead to neurobehavioral outcomes."
"While the OEHHA assessment shows active "hits" for estrogenic and glucocorticoid effects in summary Table 4.1 (pages 152–154), these activities were not discussed further."

Response 4.10: The purpose of exploring the HTS data for the food dyes in the OEHHA report was to survey the assay results as a supplemental line of evidence for any potential mechanisms by which the food dyes might interact with neurological targets. We agree that there is very limited evidence from the HTS assays for the mechanisms by which Red No. 40 might impact the nervous system. OEHHA cannot say whether there was evidence to support or dispute an association between exposure to Red No. 40 and neurobehavioral-related key events. Our work did not explore key events, but rather activity of the dyes on molecular targets. Note that we discuss potential mechanisms of Red No. 40 in Section 4.1.3 of the report; Red No. 40 inhibits estrogen receptor in vitro, probable metabolites of Red No. 40 inhibit the esterase carbonic anhydrase, and Red No. 40 reduces cell viability in neuronal progenitor cells.

The potential role of glucocorticoid (GC) receptors in nervous system is discussed in the OEHHA report (page 159) stating that "GCs and their receptors exert widespread actions in the central nervous system, ranging from the regulation of gene transcription, cellular signaling, and modulation of synaptic structure. Elevated GC levels are linked to neuronal plasticity and neurodegeneration (Vyas et al. 2016)." Further, the estrogen receptor is involved in brain development, and estrogens modulate many processes in the brain, including those influencing behavior (McCarthy, 2008; McEwen et al., 2012). As noted by the McEwan et al. (2012) estrogen Is involved in a number of processes in the brain that

> " include fine motor control, motor coordination, pain, mood regulation, cognitive function, cardiovascular regulation, neuroprotection and many others. These behaviors and brain functions involve brain areas beyond the hypothalamus, including the spinal cord, cerebellum, nigrostriatal and mesolimbic system, amygdala, hippocampus, cerebral cortex and brainstem, as well as a large array of neurotransmitter and neuromodulator systems, including cholinergic, noradrenergic, serotonergic, dopaminergic, glutamatergic, neuropeptide Y and opioidergic systems."

We have added text on the role of estrogens in the brain.

Comment 4.11: (Appendix B, page 5) "The OEHHA assessment, which did not apply filters for either data quality issues or cytotoxic interference in the determination of activity, considered Blue No. 2 active for interaction with the aryl hydrocarbon receptor, and shows multiple active "hits" across other neuro-relevant categories in summary Table 4.1 (pages 152–154). However, these active assays/categories were not discussed further in the assessment."

Response 4.11: Blue No. 2 was active in assays for the following molecular targets: aryl hydrocarbon receptor, androgen receptor, and the thyroid hormone receptor. OEHHA provided discussion on numerous neuro-relevant targets (nuclear receptors, dopaminergic and serotonergic receptors, and the glucocorticoid receptor (NR3C1) in Chapter 4 as well as in Appendix A describing the rationale for choosing HTS assays. As noted in the document, our evaluation of HTS assay results was exploratory and the purpose was to try to identify potential molecular targets relevant to neurological endpoints generally.

Although not specifically with food dyes, the aryl hydrocarbon receptor has been shown to mediate zebrafish neurogenesis and gliogenesis (Wu et al., 2019) and has been linked to the mediation of neurological activities in other studies (Wójtowicz et al., 2017; Choudhary et al., 2020). Therefore, it is reasonable to include this nuclear receptor as a potential neurological marker. Further discussion of the results of the food dyes and aryl hydrocarbon receptor has been added to Chapter 4 and Appendix A, where the methodology of assay selection is discussed.

Comment 4.12: (Appendix B, page 15-16) "The OEHHA assessment, which did not apply filters for either data quality issues or cytotoxic interference in the determination of activity, considered Yellow No. 5 active for aryl hydrocarbon receptor downregulation and showed multiple active "hits" across other neuro-relevant categories as summarized in Table 4.1 (pages 152–154). These results are not further discussed. The inclusion of the aryl hydrocarbon receptor assay(s) is related to "AOP 150" according to Table 4.1 (page 153). Assuming this refers to AOPWiki AOP #150, which is "Aryl hydrocarbon receptor activation leading to early life stage mortality, via reduced VEGF," it is not clear how this AOP is related to neurobehavioral outcomes.

Response 4.12: The notes section of Table 4.1 is not intended to be a rationale section for the inclusion of the assays or molecular targets and was never stated as so. The rationale for inclusion of assays associated with the aryl hydrocarbon receptor was stated in the Methods section of Appendix A. The association of this assay to AOP 150 (in the notes section) was identified to show alternate links to other adverse pathways and was not intended to infer an

association between AOP 150 and neurotoxicological or neurobehavioral outcomes. Finally, neurobehavioral effects are a form of neurotoxicity and mechanistic data is not available for all chemicals that have neurobehavioral effects. Thus, OEHHA explored HTS assays for any molecular targets that might be associated with neurotoxicity.

Comment 4.13: (Appendix B, page 13) "The OEHHA assessment, which did not apply filters for either data quality issues or cytotoxic interferences in the determination of activity, considered Yellow No. 6 active for cholinergic effects and shows multiple active "hits" across other neuro-relevant categories in summary Table 4.1 (pages 152–154), without further explanation."

Response 4.13: The activity of Yellow No. 6 associated with the cholinergic receptors is based on assay NVS_GPCR_hM3 (refer to Appendix D of the OEHHA Report). Please see OEHHA's response to comment #2 regarding data quality issues and cytotoxicity inferences. OEHHA agrees that because there are several flags associated with this assay result, other supporting information should also be considered.

Comment 4.14: (Appendix B, page 30-31) "…OEHHA assessment states in Section 4.3.3.4, page 159: "All dyes were active in assays targeting dopaminergic […] receptor subtypes." In contrast to the quoted statement, according to Table 4.1 on page 152 in OEHHA's report, Blue No. 1, Blue No. 2, and Green No. 3 are listed as "not tested" in assays for "dopaminergic" molecular targets. The assay data in Table 1, Appendix C of the OEHHA assessment corroborate that Blue No. 1, Blue No. 2, and Green No. 3 were not tested in assays for dopamine receptor subtypes."
"In contrast to the quoted statement, according to Table 4.1 on page 152 in the OEHHA Draft Report, Blue No. 1, Blue No. 2, and Green No. 3 are listed as "not tested" in assays for "opioid" molecular targets. The assay data in Table 1, Appendix C of the OEHHA assessment corroborate that Blue No. 1, Blue No. 2, and Green No. 3 were not tested in assays for opioid receptor subtypes.

Response 4.14: Thank you for pointing out this error. The descriptor "azo" should have appeared between "all" and "dyes". We have corrected the statement in the OEHHA report.

Comment 4.15: (Appendix B, page 30) "Regarding dopaminergic activity in HTS assays, both OEHHA and Chappell et al. (2020) show Red No. 3 to be active. This result, integrated with *in vivo* mechanistic and neurobehavioral findings, is discussed at length within this document above in the section focused on Red No. 3 data. The other colors that were tested in dopamine receptor subtype assays (Red No. 40, Yellow No. 5, and Yellow No. 6) were considered inactive when cytotoxic assay interference and data quality flags were considered. Specifically, the AC50 values for all three of these colors in the assay for DRD1 loss-of-signal is well above the respective lower bound estimate for cytotoxicity for each color, according to the ToxCast summary files: Red No. 40, AC50 = 25.30μM vs. cytotoxicity lower bound of 6.60μM; Yellow No. 5: AC50 = 21.62μM vs. cytotoxicity lower bound of 8.14μM; Yellow No. 6, AC50 = 34.40μM vs. cytotoxicity lower bound of 7.94μM. Test article purity and identity information was not available for the dopaminergic signaling HTS assays.

Response 4.15: Please refer to OEHHA's response to comment #2 regarding our view of data flags. We did not disregard the results of HTS assays based on whether or not there were any data flags including for cytotoxicity, in line with Judson et al., (2016). Chappell et al. (2020) did not cite any support for the dismissal of active assays with more than one flag for data quality issues. There is as yet no established method for discarding data based on data flags.

The comment does not provide references to support using "cytotoxic assay interference" and "data quality flags" to render an assay "inactive" (other than their own paper, Chappell et al., (2020). In many cases, Chappell et al. (2020) did not use the hit calls in the ToxCast database in their analysis, but rather used data flags to change the hit calls from active to inactive. By throwing out all assay results that had more than one data flag, the commenters have disregarded much of the available information from the HTS assays. We agree that data flags increase uncertainty and complicate interpretation. We disagree that one should therefore simply dismiss those results.
At the very best, an active hit at a concentration below the cytotoxicity limit indicates that the active hit is not likely due to processes involved in cell death (R. Judson, personal communication, April 3, 2020). However, the inverse, that active hits above the cytotoxicity limit are likely due to processes involved in cell death, cannot be applied to justify ignoring those assays. It may cast uncertainty over the result, but we do not know in fact that the result is due to processes involved in cell death. In addition, from this communication with Dr. Judson, the "cytotoxicity limit is not a hard cut-off; some of the active hits in the grey zone [above the cytotoxicity limit] are real."

Comment 4.16: (Appendix B, page 30-31) "Gentry et al. (submitted) identified inconsistent changes in in vivo measures of oxidative stress in the brain, neuronal cell damage, and neurotransmitter levels (e.g., serotonin, dopamine, and GABA), with no corresponding measures in behavioral activity evaluated in most mechanistic studies. For in vitro data, only a single serotonin receptor binding (loss of signal) HTS assay was active among otherwise inactive endpoints measured in in vitro neuro-relevant mechanistic assays (Chappell et al., 2020). Together, these results do not support biological plausibility between exposure to Yellow No. 5 and neurobehavioral effects. There was insufficient evidence supporting an association between exposure to Yellow No. 5 and neurobehavioral-related key events." "The OEHHA assessment states in Section 4.3.3.4, page 159: "Blue No. 1, Red No. 40, and both yellow dyes were also active for serotonergic receptors."

"Although Yellow No. 5 was active in a *single* assay for serotonergic activity, this loss of signal for the 5-hydroxytryptamine (serotonin) receptor 1A (Htr1a) in rat cortical membranes appears to be species-specific, because other serotonin receptor binding assays for human and guinea pig serotonin receptor genes were inactive for Yellow No. 5. Also, we disagree that the other listed colors were active in serotonergic assays. To illustrate the decision criteria that should be considered in determining activity/inactivity in such assays, see Figure 2 below [This figure can be viewed in the original comment letter]. Test article purity and identity information was not available for the serotonergic signaling HTS assays"

Response 4.16: OEHHA agrees that concordant observations across multiple active assays for one specific target can provide greater confidence in a particular effect. Given the small number of active assays for the food dyes, however, it would be surprising to observe such consistency. In fact, lack of biological concordance among related assays is commonly seen in high-throughput assays, even for chemicals with more active assays than were observed with some of the food dyes (Truong et al., 2014; Silva, et al., 2015).

The decision criteria illustrated in Figure 2 contradicts IACM's statement that the other listed colors (Blue No. 1, Red No. 40, and the yellow dyes) should not be considered active in the serotonergic assays. For each of the colors mentioned, there were at least two measurements above the baseline cutoff for activity (horizontal grey line). If the top estimate value of the model and maximum median value is both greater than or equal to the efficacy cutoff, the assay is considered active – there had to be at least one data point above the cutoff value to be active (R. Judson, personal communication, April 3, 2020). This is true for the following serotonergic assays: NVS_GPCR_h5HT7 (Blue No. 1 and Red No. 40) and NVS_GPCR_r5HT1_NonSelective (Yellow No. 5 and Yellow No. 6).

Comment 4.17: (Appendix B, page 32) "Opioid receptors were not evaluated in Chappell et al. (2020) as these were not identified as having direct relevance to neurobehavioral outcomes according to the resources reviewed within the Chappell et al. (2020) assessment (e.g., literature reviews, AOPWiki, and the Comparative Toxicogenomics Database, among others). Nevertheless, in evaluating the available evidence OEHHA considered, we noted the following: only Red No. 3 was active in any of the opioid receptor subtype assays, none of the other colors were active in opioid receptor subtype assays when the same criteria for data quality and cytotoxic assay interference as discussed above (and as used in Chappell et al., 2020) were applied. Red No. 3 was tested in two assays for opioid receptor subtypes and was active for loss of signal for the human opioid receptor, mu 1 (*OPRM1*) in a cell-free radioligand binding assay. Red No. 3 was inactive in an assay for guinea pig opioid receptor, kappa 1 (*Oprk1*). A potential relationship between opioid receptors and mechanisms of neurobehavioral outcomes was not presented by OEHHA. Nonetheless, it would be expected that a potential adverse effect would be related to activation of opioid receptor subtypes, whereas Red No. 3 was only active for *loss* of signal for OPRM1. Test article purity and identity information was not available for the HTS assays related to opioid receptor subtypes."

Response 4.17: The intent of the exploration of the HTS data, was to assess potential underlying molecular mechanisms that may be linked to neurological pathway perturbations. OEHHA disagrees that opioid receptors have no direct linkages to neurological outcomes. For example, from the Comparative Toxicogenomics Database (CTD), there are inferred associations between the opiate receptor *kappa 1* (evaluated in our set) and neurotoxicity syndromes, neural tube defects, and neurobehavioral manifestations. Other opiate receptor subtypes also have inferred associations with neurological diseases on CTD.

OEHHA did not use the same criteria approach as Chappell et al. 2020 for reasons stated in the OEHHA report (page 158). Therefore, our conclusions differ, and we stand by our analysis.

It is unclear how the cytotoxicity assay interference was incorporated for the opioid receptor subtype assays evaluated in Chappell et al. 2020. Red No. 3 was active in the assay NVS_GPCR_hOpiate_mu and Red No. 40 was active for assay NVS_GPCR_hOpiate_D1 (both active assays had zero flags for efficacy). Both assays are associated with a loss of signal as an indicator of binding activity therefore they should be considered for further investigation. The activity of the yellow dyes with the opiate receptor subtype assays do have flags (please see OEHHA's response to comment #2 regarding flags) and therefore should still be considered but potentially viewed as less reliable.

Comment 4.18: (Appendix B, page 13) "OEHHA included assays with relatively broader (i.e., indirect, unclear, and/or non-specific) potential relationships to neurobehavioral adverse outcomes, such as estrogen receptor (E.R.), androgen receptor (A.R.) antagonism/agonism, and inflammation in non-neuronal cells, respectively. These should not be included as inflammatory response in other cell systems is too general to draw conclusions regarding the ability of the color tested to induce inflammation in the brain *in vivo*, and a link between androgenic or estrogenic changes and neurobehavioral outcomes has not been identified to be a key event in any of the pathways identified to date for neurobehavioral adverse effects as noted in Chappell, et al. (2020)."

Response 4.18: Please see OEHHA's response to comment #1 regarding evaluation of inflammation in non-neuronal cells. There is support for food dye activity in association with the estrogen receptor (Dees et al., 1997; Axon et al., 2012). Although there is less literature available on food dye activity with the androgen receptor, we felt this was also a valid target to explore based on 1) the work of Mathieu-Denoncourt et al., 2014 which explore the activity of the azo dye Yellow 7, albeit not a dye in our set; 2) the well-known fact that ARs and ERs often act and regulate in concert; 3) the broad influence of estrogens and androgens on brain function. The intent of the current work is not to draw conclusions on whether the dyes demonstrate ER or AR antagonism/agonism or induce inflammation in the brain *in vivo*, but rather to address potential concordance with literature, and to explore potential pathways of perturbation that may eventually be linked to downstream neurological effects.

As for the relationship between estrogenic, anti-estrogenic effects on behavior, there is substantial neuroendocrinology literature demonstrating a number of pathways by which estrogen regulates many neuronal processes, including behavior (see for example McEwen et al., 2012). Estradiol actively modulates many processes, both genomic (e.g., modulating gene expression) and non-genomic (e.g., interactions with cell membrane receptors resulting in direct, rapid effects on synaptic function involving cells with varied neurotransmitter systems) in the brain. As well, testosterone is metabolized in the brain to estradiol. Further, both nuclear and non-nuclear androgen receptors are present in cells in many parts of the brain. The capacity for estrogens and androgens to modulate brain activity are the reasons OEHHA opted to include assays related to endocrine pathways in assessing potential mechanistic pathways for effects of the food dyes on the nervous system.

Comment 4.19: (Appendix B, page 32-33) "OEHHA assessment Section 5.4, page 172: "All the food dyes were active for the androgen assays tested. The dyes, except for Blue No. 2 and the yellow dyes, were active for the receptor-based antagonist assays for the estrogen receptor, potentially indicative of antagonism for this receptor."

There is no justification or rationale for the inclusion of estrogenic and androgenic assay data provided in the OEHHA assessment, nor discussion regarding the potential effects of ER or AR activity on neurobehavioral outcomes. Estrogenic and androgenic effects were not considered in Chappell et al. (2020), as estrogenic and androgenic effects were not identified as having direct relevance to neurobehavioral outcomes according to the resources reviewed within the Chappell et al. (2020) assessment (e.g., literature reviews, AOPWiki, and the Comparative Toxicogenomics Database, among others) for mode of action and/or underlying mechanisms of neurobehavioral outcomes."

"Nevertheless, upon evaluation of the ER and AR ToxCast Pathway Models, which integrate a battery of HTS assays that represent events across the ER and AR pathways, inactivity was predicted for agonist, antagonist, and receptor binding for all seven colors, with the exception of a weak AR antagonist activity for Blue No. 1 and Green No. 3. These computational models were developed for use by the Endocrine Disruptor Screening Program (EDSP) and provide results that can discriminate bioactivity from assay-specific interference and cytotoxicity in the integration of assay data from 18 ER assays or 11 AR assays. Blue No. 1 and Green No. 3 had an area under the curve (AUC) value, which is the model output, of 0.107 and 0.209, respectively, for AR antagonism, indicating weak antagonist AR activity (an AUC value of 0.1 corresponds to activity at ~100 µM by this model)."

Response 4.19: Please see OEHHA's responses to comments #10 and #18 above. There is substantial literature demonstrating the effects of estrogens and androgens on many brain functions including those involved in behavior. For our exploratory analysis of the HTS assay results, we chose to include these important receptors precisely because of the broad role of estrogens and androgens on brain function including behavior. IACM did not provide details or a citation for their analysis of the HTS data using the EDSP computational model.

Comment 4.20: (Appendix B, page 26) "We agree that, based upon a single assay, Blue No. 1 and Green No. 3 suggest the potential for thyroid antagonistic activity (as reported in Chappell et al., 2020). However, such a hypothesis would have to be tested further.
"We disagree that the data for Blue No. 2, Red No. 3, and Red No. 40 should be considered active. Data for these colors in this single assay evaluating thyroid receptor (TR) antagonism are either unreliable or inactive because of serious data quality issues (Blue No. 2) or due to significant loss of cell viability in that assay (Red No. 3 and Red No. 40). Test article purity was sufficient and identity was confirmed for Red No. 3 and Red No. 40, while the test article was considered impure by nuclear magnetic resonance (NMR) for Blue No. 2, and no information was available for identity."

Response 4.20: The source for the test article purity for the food dyes as referenced by IACM is unclear. Per the Chemistry Dashboard, one of the Tox21 identifications for Blue No. 2 has a purity > 90% for T4.

All four potential data flags are associated with the TOX21_TR_LUC_GH3_Antagonist assay hit for Blue No. 2 and therefore reliability for this assay is considerably less. There are no data flags for this assay for Red No.3, and only the "less than 50% efficacy" flag associated with Red No.

40. However, in considering cytotoxicity, because the associated cell viability assay (TOX21_TR_LUC_GH3_Antagonist_viability) was also from the Tox21 platform and had no data flags, it is more likely that the AC50s observed for the Red No. 3 and Red No. 40 activities are influenced by cytotoxicity.
The OEHHA report (section 4.3.3.4) will be revised to show that less reliability is associated on the interaction of these three dyes with thyroid receptor antagonism. However, the data for Blue No. 2, Red No. 3, and Red No. 40 should still be considered active, understanding that the reliability is uncertain.

Comment 4.21: (Appendix B, page 23) OEHHA assessment Table 4.1, page 154: "Four dyes [Blue No. 1, Green No. 3, Red No. 3, Red No. 40] associated with downregulation; targeting the loss of signal of TPO activity. The assay is associated with AOP 42" [Inhibition of Thyroperoxidase and Subsequent Adverse Neurodevelopmental Outcomes in Mammals]."

"We do not agree that the TPO assay for these four colors should be considered active. For all four colors, the matched viability assays that are specifically associated with the TPO assay (as reported in a 2016 publication from US EPA scientists within the NCCT/CCTE) are also active, and have AC50 values below that of the AC50 for the TPO inhibition assays, as shown in Table 1 below. These results indicate that these colors cause a loss-of-signal for TPO enzyme activity only at concentrations at which significant cell death also occurs *in vitro*. Test article purity and identity information was not available for these assays."

Response 4.21: Please see OEHHA's response to comment #2 regarding cytotoxicity as a cutoff. It should be noted that only Red No. 40 had a single flag for efficacy associated with the TPO assay; the other three dyes had no flags associated with their assay activity. The 2016 publication (Friedman et al., 2016) cited the utilization of NCCT_HEK293T_CellTiterGlo as an indicator of possible sources of nonspecific assay signals. The publication also stated selection of chemicals where the activity of the TPO assay were at concentrations lower than those associated with the cytotoxicity assay. Friedman et al. states that "Reactive chemicals tested at concentrations exceeding cellular toxicity limits and nonspecific enzyme inhibition caused by detergents, salts, or cross-reactive compounds can potentially confound the interpretation of activity in the AUR-TPO assay."

OEHHA notes this and will take this into consideration in the interpretation of the results associated with TPO activity. At the very best, an active hit at a concentration below the cytotoxicity limit indicates that the active hit is <u>not likely</u> due to processes involved in cell death (R. Judson, personal communication, April 3, 2020). However, the inverse, that active hits above the cytotoxicity limit are likely due to processes involved in cell death, is not necessarily correct and should not be applied to justify ignoring those assays. It may cast uncertainty in interpreting the results, but because it is not known for certain that the result is due to processes involved in cell death, we should not automatically exclude them.

Response to Comments from Consumer Products Healthcare Association

Comment 1: (page 1) "Prior to their use in food, drugs, or cosmetics, the Food and Drug Administration (FDA) must approve color additives. FDA also maintains and regularly monitors postmarket surveillance databases to which consumers, health professionals and industry can submit adverse events believed to be related to color additives (as well as other products). Color additives have been safely used in a wide variety of consumer products for decades and given the well-established role of FDA in regulating these color additives, consideration must be given to the potential for consumer confusion should OEHHA suggest additional regulatory action."

Response 1: We agree that FDA is responsible for regulating food dyes. Regulation beyond what FDA is doing is not in OEHHA's purview.

Comment 2: (page 2) "The OEHHA report analyzes the available data on the possible effects of synthetic food dyes (*i.e.*, color additives) on neurobehavior in children, including epidemiologic and animal studies. However, a weight of evidence analysis reveals that a causal association between exposure to food dyes and adverse neurobehavioral effects has not been demonstrated. FDA and other scientific expert bodies have been examining this issue since 1982 and have yet to find strong, consistent evidence linking color additive intake with adverse neurobehavioral effects including Attention Deficit Hyperactivity Disorder (ADHD). A series of robust reviews of the available scientific evidence have routinely found that a causal relationship between color additive intake and adverse neurobehavioral effects does not exist."

Response 2: In 2011, US FDA asked its Food Advisory Committee (FAC) to consider available relevant data on the possible association between consumption of synthetic color additives in food and hyperactivity in children, and to advise FDA as to what action, if any, is warranted to ensure consumer safety. US FDA concluded that most children have no adverse effects when consuming foods containing color additives, but some evidence suggests that certain children may be sensitive to them.

The FDA states on their web page "*The FDA has reviewed and will continue to examine the effects of color additives on children's behavior. The totality of scientific evidence indicates that most children have no adverse effects when consuming foods containing color additives, but some evidence suggests that certain children may be sensitive to them…."*

As we describe in the Introduction, OEHHA did not limit the review to the question of effects on children diagnosed with ADHD or other behavioral disorders. Rather, OEHHA evaluated the literature to determine whether there might be any effects on behavior of the FD&C batch-certified synthetic food dyes in children in the general population with or without a diagnosis of ADHD. We did not focus solely on effects related to activity and attention, but evaluated the literature for effects on other neurobehavioral impacts as well. In addition, OEHHA evaluated the animal toxicology literature relevant to neurological endpoints; these studies were not emphasized in the 2011 US FDA review.

As can be seen in Tables 3.9 and 3.10 in the OEHHA report, there were a number of animal studies with statistically significant changes in behavior following administration of individual food dyes as well as mixtures of food dyes. The other authoritative bodies did not adequately weight these effects. Several well-conducted studies were published after the reviews of USFDA and in some cases after the updates from JECFA. Additional studies have been published since then, both in vivo and in vitro. Finally, we reviewed newer data relevant to mechanisms of action of potential neurobehavioral or neurotoxic effects of the food dyes. The authoritative bodies cited in the comment did not review and integrate all the mechanism,

animal toxicology and human studies on the topic of neurotoxicity.

In short, OEHHA did a more thorough review of more scientific information than any of the international bodies noted in the comment. It should be noted that the EU requires that foods with azo dyes be labeled with a warning "may have an adverse effect on activity and attention in children". (https://www.europarl.europa.eu/sides/getDoc.do?language=EN&type=IM-PRESS&reference=20080707IPR33563)

Many of the animal studies cited by OEHHA found effects at the level of the NOAELs from these old studies used as a basis for the ADI, and in a few cases at the level of the ADI. IACM's own contractors, while dismissive of many studies for what we consider inappropriate reasons (see comments below to specific studies), judged most of the individual dye studies OEHHA cites as "reliable without restrictions" or "reliable with restrictions" using the ToxR Tool and note a number of them scored high in their application of the ARRIVE guidelines (see Appendix A, Table A1).

Most importantly, there is substantial evidence from human studies that the food dyes impact behavior in some children. As noted above, FDA acknowledges this. OEHHA presents a thorough systematic review of the literature in chapter 2 of the OEHHA report, including a comprehensive study quality evaluation. The clinical trial studies in children, where children were first given a diet free of food dyes and then provided food dye challenge, and a meta-analysis published in 2012 (Nigg et al, 2012) support an effect of food dyes on children's behavior. One study (Stevenson et al., 2010) found that certain polymorphisms for histamine breakdown (histamine N-methyl transferase gene) are associated with an increased risk of children responding to dyes in the dye mixture studies. Interestingly, recent scientific studies have shown that histamine is a neurotransmitter and histaminergic neurons interact with many other structures in the brain including those involved in behavior (Yoshikawa et al, 2019; Scammel et al., 2019). OEHHA's report states (in Section 2.8), after a thorough evaluation of the human clinical trials, "Based on the extent of the positive findings reported, and the fact that we could not convincingly or consistently attribute these positive findings to errors in study design or other bias, we conclude that the current human epidemiologic evidence supports a relationship between food dye exposure and adverse behavioral outcomes in some children, both with and without pre-existing behavioral disorders."

Comment 3: (CHPA letter, page 2) "Studies used to support an association between intake of color additives and adverse neurobehavioral effects suffer from a series of limitations including examination of mixtures of colors, unverified validity of behavior scores, small sample size and lack of a dose-response relationship."

Response 3: Please see our response to the other comments below (as well as our responses to the comments from IACM) regarding these issues. All of these were already considered in OEHHA's report.

Comment 4: **(CHPA letter, page 6)** "Results from a 2012 meta-analysis (1), which covered largely the same studies contained in the OEHHA report, suggested that only a weak conclusion could be drawn due to the extent of heterogeneity in study results. The review by OEHHA includes all 24 studies in Nigg et al plus 4 additional studies (2 that were pilot studies and 2 other studies with samples sizes of less than 3). These additional studies would not likely have changed the assessment reported by Nigg et al. (2012)."

"Nigg et al. stratified their analyses by who conducted the neurobehavioral assessment (parent or teacher/observer) because this was a known source of heterogeneity. The combined estimate when parents did the testing was 0.18 (95% CI 0.08–0.29), a moderate and statistically significant association. OEHHA also noted that assessments by parents were more likely to report a statistically significant effect. The combined estimate when a teacher (or observer) did the testing was 0.07 (95% CI -0.03–0.18), a much smaller, nonsignificant association."

Response 4: OEHHA's evaluations of each of the relevant studies, and of the literature as a whole, went far beyond those presented by Nigg et al. (2012). For example, while Nigg et al. (2012) included essentially no discussion of the relevant laboratory animal studies or of biologic plausibility, OEHHA's report includes a 25 page review and discussion of these animal studies and a 14 page review and discussion of the toxicokinetic and mechanistic data. With regards to the human epidemiologic data, OEHHA's report includes a much more detailed description of each study (Table 2.1), a much more thorough and transparent assessment of each studies major strengths and weaknesses (Table 2.3a), and a much more extensive discussion of causal inference. For example, while the Nigg et al. (2012) publication provided information on nine characteristics for each study, OEHHA's tables included information on 41 different study characteristics. And, while the ratings of study quality in the Nigg et al. (2012) meta-analysis were based on only 5-6 variables, OEHHA's quality assessment included 20 different variables (page 31). Overall, while the Nigg et al. (2012) meta-analysis is a very useful part of OEHHA's evaluation, it is only a part. OEHHA's evaluations of the literature as a whole and on causal inference are many times more extensive than those provided by Nigg et al. (2012).

While parent scoring may have been noted in a portion of the Nigg analysis as providing a stronger association with behavior, our review of the Nigg meta-analysis included the following regarding their outcomes grouped as tests of attention: *Three neuropsychologists sorted the tasks from all studies according to the cognitive abilities that they assessed and agreed 100% on those that measured attention. Examination of these psychometric tests of attention, in many ways the more relevant metrics, yielded a consistent higher effect size, 0.27 (p= 0.007). When limiting the analysis to studies that only included FDA dyes, the effect was 0.34 but with the caveat that this included only 5 studies with a total of 68 participants. This statistically significant higher effect size using measures of attention is important since this metric is an objective measurement from experimenter table-top administration of tests. These tests of attention avoid some limitations of other measurements obtained from parents or teachers, such as rater bias, cultural effects, or stratification of studies. The results remained significant after consideration of possible publication bias and were consistent in direction with those from parent and teacher report*. Thus, overall associations have been found across studies beyond parent report.

Comment 5: (CHPA letter, page 6) "Nigg et al also prespecified other subgroup analyses of interest, all stratified by assessor. A subset of studies examined only FDA-approved colors. The combined estimated for parent assessed studies for FDA approved colors was .13 (95% CI -0.04–0.30), which was small and not statistically significant. The corresponding combined estimate for teacher/observer assessors was .12 (95% CI -0.01–0.35), another nonsignificant result.

"Another source of variability was whether the outcome measure was validated, which is considered to be of higher quality and more reproducible. When the analysis was confined to studies with a validated outcome, the parent assessed result was 0.13 (95% CI 0.00–0.25),

which is small and of borderline significance. For teacher/observer assessors, the result was small and not statistically significant (0.10; 95% CI: -0.01–0.21)."

Response 5: Overall conclusions should not be based solely on statistical significance. The reasons for this include the fact that clinical and statistical significance are not always correlated, the mostly arbitrary nature of the term and definition of "statistical significance", the fact that small effect sizes can have important population effects, and the fact that study quality and sample size are not always well correlated. These issues are described in further detail elsewhere (Hill, 1965; Pelsser et al., 2017; Amrhein et al., 2019). As mentioned previously, the results of the Nigg et al. (2012) meta-analysis cited by the commenter represent overall average effects. They do not account for the fact that some children will have greater responses than others. In addition, while the effect sizes cited by the commenter may seem small in any given "average" individual, small effect sizes like these can have enormous impacts on a population-wide basis.

Comment 6: (CHPA letter, page 6) "Many of the reviewed studies (as presented in Table 2.1 in the OEHHA report) had small sample sizes (nearly 60% of the studies included 20 or fewer children), inadequate definition of the effect size, gaps in the description of the cohort selection/recruitment process, and/or a lack of effect in the measured outcome. Indeed, OEHHA notes in their report that studies with larger numbers of participants and studies involving higher doses showed "*weak and inconsistent*" results. The smallest sample in the Nigg et al review was 4 and the largest was 277."

Response 6: Because sample size does not always correlate with study quality or validity, it is usually inappropriate to use sample size as the only reason for excluding studies or as the sole basis for making conclusions. Although some of the studies reviewed by OEHHA may have had an "inadequate definition of effect size", many did not (Table 2.1). OEHHA already notes that many of the studies we reviewed likely involved convenience samples (page 32). But, OEHHA has also already noted, that while this could have affected the generalizability of some study results, it would not necessarily affect the ability of these studies to provide useful information on causality (page 44).

Comment 7: (CHPA letter, page 6-7) "A dose-response relationship was identified in one of three studies (Rowe KS and Rowe KJ 1994, Synthetic food coloring and behavior: A dose response effect in a double-blind, placebo-controlled, repeated-measures study. J Pediatr 125:691-698) in which this information was available, but the effect was weak. Further, the authors noted that reactors were not well defined and that the dose-response was not consistent in the non-reactors. Nigg et al. also assessed dose response in a meta regression analysis separately for parents and teacher/observers. For both analyses the association was essentially zero."

Response 7: While dose-response is one of the criteria commonly used to evaluate causal inference, it is not a sine qua non (Bradford Hill, 1965). Rowe and Rowe did in fact see a dose-response for tartrazine (Yellow No. 5) exposure in children who reacted to the dye. The behavioral scores were highly statistically significantly different than the non-reactor group. The fact that the nonreactors did not show a dose-response is irrelevant. The comment that the reactors were not well defined by the authors is misleading. The reactors are defined by the behavioral changes induced by exposure to tartrazine. Perhaps the basis of the comment is that the underlying reasons why these children reacted to the dye may not be known.

Regardless, Nigg et al. (2012) did identify an association between artificial food dye dose and behavioral response in their meta-regression (slope (β) = 0.002, p = 0.034). No details were given on how dose was entered into this regression. As such, it is not clear how the commenter came to the conclusion that this association is "essentially zero". In addition, given that the studies included in the Nigg et al. (2012) meta-analysis were fairly heterogeneous in terms of the dyes used, the use of mixtures, the duration of exposure, outcome assessment, and a number of other factors, it is not clear that the meta-regression used by Nigg et al. (2012) was an appropriate way to evaluate dose-response.

Comment 8: (page 7) "The study conducted by Bateman et al. (2004) is one of the few reported clinical trials that seems to support a possible effect of color additives on behavior, having a large sample size and showing statistically significant effects for both the improvement of hyperactivity while on the elimination diet and the reappearance of hyperactivity upon re-challenge. However, increases in hyperactivity were only observed upon parental evaluation not following evaluation by professional psychological examinations."

Response 8: We acknowledge the heterogeneity of study designs and outcome measures and findings. Individual studies have had differences in which metrics found significant associations with exposure. We disagree that Bateman was "one of the few" studies supporting a possible association between food dyes and behavioral responses. We did not limit our considerations to children with the diagnosis of ADHD but rather to any neurobehavioral impacts related to dye exposure. While in some instances parent score is noted to be higher than teacher or observer scoring, the Nigg meta-analysis noted psychometric tests of attention, in many ways highly relevant metrics, yielded a consistent higher effect size, 0.27 (p= 0.007). These tests of attention avoid some limitations of other measurements obtained from parents or teachers, such as rater bias, cultural effects, or stratification of studies. Please see our response to comment #4.

Comment 9: "In their report, OEHHA presents exposure assessments for certain OTC medications indicated for children (pain reliever/fever reducer syrups; cold, cough and allergy syrups as well as dietary supplements). The methodology employed to measure color additive levels in these products is described in greater detail in a recent publication. Based on results presented in the report, OEHHA notes that exposures above the ADI (for FD&C Red No. 40) would occur with certain brands of children's pain reliever/fever reducer syrups or cold, cough and allergy syrups if the maximum labeled daily dose were to be taken. These values were noted to be greater than 10-40 times higher compared to estimated intakes from food (based on data from NHANES)."

"In a recent discussion held between CHPA members, OEHHA and UC Davis regarding the methodology described in the Lehmkuhler et al., 2020 paper and referenced in the OEHHA report[28], an error was identified in the calculations used to estimate color additive exposures from children's cough/cold and allergy syrups. This error led to erroneous values being reported for the children's pain reliever/fever reducer syrups and cold, cough and allergy syrups noted in both the Lehmkuhler et al. paper as well as the OEHHA report."

"CHPA asks that OEHHA issue a revised report, with a corresponding request for comments, providing corrected estimated exposure values as well as additional detail of the methodology utilized to measure FD&C colors in OTC medicines, dietary supplements and foods as well as corrections to the following:

- Erroneous values noted for estimated exposure (in mg/kg/day) to FD&C Red No. 40 and Blue No. 1 (off by a factor of 100)
- Statements regarding hazard indices (margins by which estimated exposures exceed an ADI)
- Incorrect statements in the report related to exposures for FD&C Red No. 40 being above the Acceptable Daily Intake (ADI).
- We also note that OTC medications studied in this report are not indicated for chronic use."

Response 9: In response to the reviewers' comments, we confirmed the inaccuracies in our initial color additive exposure estimates for children from cough/cold/allergy and pain relieve/fever reducer syrups, and have made the changes detailed below to the text and tables of the report. With one exception, the corrected values are all 100-times lower than initially reported and none of the estimated exposure values exceed the US FDA or JEFCA ADIs. We believe we have addressed each of the reviewers' comments and concerns.

Page 224 Section 6.8.1

Section 6.8.1 was updated to contain the corrected values. We revised the text as follows:

> Original: "…The estimated FD&C Red No. 40 exposures ranged from 2.90 to 3.15 mg/kg/day for 1 dose/day and 10.1 to 12.9 mg/kg/day for 4 doses/day. The estimated FD&C Blue No. 1 exposures ranged from 2.90 to 3.15 mg/kg/day for 1 dose/day and 1.80 to 2.30 mg/kg/day for 4 doses/day…"

> Revised: "…The estimated FD&C Red No. 40 exposures ranged from **0.029 to 0.032 mg/kg/day** for 1 dose/day and **0.10 to 0.13 mg/kg/day** for 4 doses/day. The estimated FD&C Blue No. 1 exposures ranged from **0.005 to 0.006 mg/kg/day** for 1 dose/day and **0.018 to 0.023 mg/kg/day** for 4 doses/day…"

Page 225 Table 6.17 - All values reported in the table were updated with the corrected values.

Page 226 Section 6.8.2 - Section 6.8.2 was updated to contain the corrected values. We revised the values as follows:

> Original: "…The estimated FD&C Red No. 40 exposures from Brand 5 ranged from 2.8 to 3.7 mg/kg/day for 1 dose/day and 16.9 to 22.1 mg/kg/day for the maximum recommended dose of 6 doses/day. The estimated FD&C Blue No. 1 exposures from Brand 5 ranged from 0.4 to 0.5 mg/kg/day for 1 dose/day and 2.2 to 2.9 mg/kg/day for 6 doses/day (Table 6.18)."

> Revised: "…The estimated FD&C Red No. 40 exposures from Brand 5 ranged from **0.028 to 0.037 mg/kg/day** for 1 dose/day and **0.17 to 0.22 mg/kg/day** for the maximum recommended dose of 6 doses/day. The estimated FD&C Blue No. 1 exposure from Brand 5 ranged from **0.004 to 0.005 mg/kg/day** for 1 dose/day and **0.022 to 0.029 mg/kg/day** for 6 doses/day (Table 6.18)."

Page 226 Table 6.18 - Values in the table were updated with the corrected values.

Page 227 Section 6.8.3 -Section 6.8.3: Paragraph 1 was revised to state that Brand 3 was the gummy vitamin brand with the highest estimated exposure for Red No. 40. We changed the text as follows:

> Original: "…The highest estimate was for Red No. 40 from Brand 1 vitamins (Red, Orange and Purple)."
>
> Revised: "…The highest estimate was for Red No. 40 from **Brand 3 vitamins (Red, Yellow and Green)."**

Page 230 Section 6.8.6 Paragraph 1 was revised to state that exposure estimates based on the maximum daily dosages did not exceed the US FDA and JECFA ADI of 7 mg/kg/day for Red No. 40. The changes were made as follows:

> Original: "…If a child were to take several doses of some brands during a single day, their intake of Red No. 40 might exceed the FDA and JECFA ADI of 7 mg/kg/day (Tables 6.17-6.18)."
>
> Revised: "…If a child were to take several doses of some brands during a single day, their intake of Red No. 40 **would not exceed** the US FDA and JECFA ADI of 7 mg/kg/day (Tables 6.17-6.18)."

Page 230 Section 6.8.6 Paragraph 2 - Section 6.8.6: Paragraph 2 was changed to present the corrected values. The text was also edited to state that exposure estimates based on the maximum daily dosages did not exceed the FDA and JECFA ADI of 7 mg/kg/day for Red No. 40. The changes were made as follows:

> Original: "Children's FD&C Red No. 40 intake estimates based on the recommended maximum daily dosages of Brand 2, grape-flavored pain reliever/fever reducer syrup, exceeded the US FDA and JECFA ADI (7 mg/kg/day) (Table 6.17; see Risk Characterization, Chapter 7). The estimated daily intakes among children 2-10 years old taking the maximum daily recommended dosage of Brand 2 syrup ranged from 10.1 to 12.9 mg/kg/day. These estimated daily FD&C Red No. 40 child intakes from Brand 2 pain reliever/fever reducer syrup were >10 times higher than mean intake estimates (typical-exposure scenario) for children 2-10 years old based on the NHANES food consumption data (Table 6.10)."
>
> Revised: "Children's FD&C Red No. 40 intake estimates based on the recommended maximum daily dosages of Brand 2, grape flavored pain reliever/fever reducer syrup, **which had the highest exposure estimates**, **did not exceed** the US FDA and JECFA ADI (7 mg/kg/day) (Table 6.17; see Risk Characterization, Chapter 7). The estimated daily intake among children 2-10 years old taking the recommended maximum daily dosage of Brand 2 syrup ranged from **0.10 to 0.13 mg/kg/day.** These estimated daily FD&C Red No. 40 child intakes from Brand 2 pain reliever/fever reducer syrup were **43% to 56% of the mean intake estimates (**typical-exposure scenario) for children 2-10 years old based on the NHANES food consumption data (Table 6.10)."

Page 230 Section 6.8.6 Paragraph 3 was revised to reflect changes in our calculation. The text was changed to state that intake estimates based on the maximum recommended daily

dosages did not exceed the US FDA and JECFA ADI for Red No. 40. The changes were made as follows:

Original: "Children's FD&C Red No. 40 intake estimates based on the recommended maximum daily doses of Brand 4, Very Berry and Brand 5, grape-flavored cold & cough syrups, also exceeded the US FDA and JECFA ADI (7 mg/kg/day) (Table 7.18; see Risk Characterization, Chapter 7). The estimated daily intakes among children 6-<12 years and 12-16 years taking the maximum daily recommended dose of Brand 4 syrup were 8.76 mg/kg/day and 9.81 mg/kg/day, respectively. These estimated daily FD&C Red No. 40 intakes from Brand 4 cold & cough syrup were >20 times higher than mean intake estimates (typical-exposure scenario) for children 12-16 years based on the NHANES food consumption data (Table 6.10)."

Revised: "Children's FD&C Red No. 40 intake estimates based on the recommended maximum daily doses of Brand 4, Very Berry and Brand 5, grape-flavored cold & cough syrups, also **did not exceed** the US FDA and JECFA ADI (7 mg/kg/day) (Table 7.18; see Risk Characterization, Chapter 7). The estimated daily intakes among children 6-<12 years and 12-16 years taking the recommended maximum daily dose of Brand 4 syrup were **0.088 mg/kg/day and 0.098 mg/kg/day,** respectively. These estimated daily FD&C Red No. 40 intakes from Brand 4 cold & cough syrup were **61% of the mean intake estimates** (typical-exposure scenario) for children 12-16 years based on the NHANES food consumption data (Table 6.10)."

Page 230 Section 6.8.6 Paragraph 4 was revised to present the corrected values. The changes were made as follows:

Original: "The estimated daily FD&C Red No. 40 intake among children 6-<12 years and 12-16 years taking the maximum daily recommended dose of Brand 5, Grape syrup, were 19.8 mg/kg/day and 22.1 mg/kg/day, respectively (Table 6.18). These estimated daily FD&C Red No. 40 intakes from Brand 5 cold & cough syrup were >40 times higher than mean intake estimates (typical-exposure scenario) for children 12-16 years based on the NHANES food consumption data (Table 6.10)."

Revised: "The estimated daily FD&C Red No. 40 intakes among children 6-<12 years and 12-16 years taking the maximum daily recommended dose of Brand 5, Grape Syrup, were **0.198 mg/kg/day and 0.221 mg/kg/day,** respectively (Table 6.18)**.** These estimated daily FD&C Red No. 40 intakes from Brand 5 cold & cough syrup were **138% of the mean** intake estimates (typical-exposure scenario) for children 12-16 years based on the NHANES food consumption data (Table 6.10)."

Page 269 Section 7.4.2 Paragraph 1 was revised to state that the exposure estimates based on the maximum daily dosages did not exceed the US FDA and JECFA ADIs. The changes were made as follows:

Original: "…If a child is treated with an over-the-counter fever reducer or pain reliever, the food dye exposures may exceed the US FDA and JECFA ADIs for some food dyes. These exposures would be in addition to any food dye exposures that day from foods."

Revised: "…If a child is treated with an over-the-counter fever reducer or pain reliever, the food dye exposures **would not exceed** the US FDA and JECFA ADIs. These exposures would be in addition to any food dye exposures that day from foods."

Page 269 Section 7.4.2 Paragraph 2 was updated to present the corrected values. The text was also revised to state that the exposure estimates based on the maximum daily dosages did not exceed the US FDA and JECFA ADI for FD&C Red No. 40. The changes were made as follows:

Original: "The Red No. 40 exposure estimates are presented in Chapter 6, Section 8. For four doses/day, one of the brands tested, Brand 2 grape-flavored pain reliever/fever reducer, exceeded the US FDA and JECFA ADI (7 mg/kg/day). The hazard indices ranged from 1.4 for children 6 to 8 years old to 1.7 for children 9 to 10 years old. No other exposure estimates from the OTC medicines tested exceed the US FDA and JECFA ADI."

Revised: "The FD&C Red No. 40 and Blue No. 1 exposure estimates **from children's pain reliever/fever reducer syrups are presented in Chapter 6, Section 8. None of the exposure estimates** exceeded the US FDA and JECFA ADIs. The hazard indices **for Brand 2 grape-flavored pain reliever/fever reducer, which had the highest exposure estimates for Red No. 40, were less than 0.02** for children 2 to 10 years old."

Page 269 Section 7.4.2 Paragraph 3 was updated to present the corrected values. The text was also revised to state that the exposure estimates based on the maximum daily dosages did not exceed the US FDA and JECFA ADI for Red No. 40. The changes were made as follows:

Original: "…The Red No. 40 exposure estimates from the maximum recommended dose of the Brand 5 grape syrup exceeded the US FDA and JECFA ADI (7 mg/kg/day). The hazard indices ranged from 2.4 for children 4 to 6 years old to 3.2 for children 12 to 16 years old. The Red No. 40 exposure estimates from the maximum recommended doses for another brand also exceeded the US FDA and JECFA ADIs, with Hazard Indices of 1.25 for 6 <12 year olds and 1.4 for 12 to 16 year olds. Thus, on a day where this medication was given, the child's exposure to Red No. 40 would exceed the dye's ADI from the medication alone. Note that some children with allergies may receive this medication chronically."

Revised: "…**None of t**he Red No. 40 exposure estimates based on the recommended maximum daily dose **of cold, cough and allergy syrups** exceeded the US FDA and JECFA ADI (7 mg/kg/day). **The hazard indices for Brand 5 grape-flavored syrup, which had the highest exposure estimates in this category, were less than 0.03 for children 4 to <12 years old and less than 0.04 for children 12 to 16 years old.** Thus, on a day where this medication is given, the child's exposure to Red No. 40 would **not** exceed the dye's ADI from the medication alone. Note that some children with allergies may receive this medication chronically."

In addition, we revised the Executive Summary and Chapter 8 sections summarizing the exposure assessment and the risk characterization to reflect the corrected information.

Comment 10: OEHHA's review of the data available from high throughput screening assays (ToxCast) measuring events possibly involved in neurodevelopment noted a relatively high number of active assay hits compared to a very low number of hits observed in a recent publication. As described in the Chappell et al. 2020 paper, of critical importance is obtaining and accounting for information on chemical purity and stability as well as issues, such as

cytotoxicity, which can potentially interfere with assay results. Indeed, OEHHA excluded a number of high throughput screening assays based on a lack of chemical quality information or issues associated with analytical results in their 2019 report 'Evidence on the Carcinogenicity of Acetaminophen'."

"Importantly, OEHHA did recognize the limitations associated with these results, noting that "…*activity for the food dyes ranged widely making it difficult to make strong correlations between what was observed, and adverse effects or mechanisms that have been reported in the literature*". Currently, regulatory agencies such as the Environmental Protection Agency and the Center for Computational Toxicology and Exposure do not use high throughput screening assays for either hazard or risk assessment owing to a disconnect between what is measured in a high throughput screening assay and the outcome of concern."

Response 10: We agree that data flags signal increased uncertainty and complicate interpretation. We disagree that one should simply dismiss those results and we incorporated all the results into our analysis regardless of data flags. Judson et al (2016) states "However, given quantitative uncertainties in AC50 values, and our incomplete understanding of the biological linkages between cell-stress, cytotoxicity, and specific activity, a hard filter will not always be appropriate." and in personal communication with the authors of that study (R. Judson, personal communication, April 3, 2020), flags and cytotoxicity limits should not be used automatically as hard filters, but should be considered when evaluating and interpreting HTS results regarding the degree of uncertainty that may be associated with the HTS results. In addition, flags are subject to change as the ToxCast data analysis pipeline evolves. Our results were extracted from the iCSS Dashboard in 2019. Since then, much of the data has been re-pipelined and new data has also been incorporated. This could have resulted in changes to the parameters that contribute to defining the hit call, which most often affects borderline hits and flags associated with them. Thus, it stands to reason that flags may change in different and future versions of ToxCast.

Further, there is not an established standardized way of incorporating cytotoxicity when interpreting ToxCast HTS data. Viability assays were given limited consideration in our evaluation due to a number of factors including 1) the variable number of cytotoxicity assays for each chemical; 2) appropriateness of utilizing Tox21 cytotoxicity assays in application to ToxCast assays outside the Tox21 platform (potentially important differences in the methods across platforms which would affect cell viability and non-concurrent measurements); 3) many cytotoxicity assays have flags of efficacy, which by the same rules as used by ToxStrategies, would render them less reliable. At the very best, an active hit at a concentration below the cytotoxicity limit indicates that the active hit is <u>not likely</u> due to processes involved in cell death (R. Judson, personal communication, April 3, 2020). However, the inverse, that active hits above the cytotoxicity limit are likely due to processes involved in cell death, is not necessarily accurate and cannot be applied to justify ignoring those assays. It may cast uncertainty over the result, but we do not know in fact that the result is due to processes involved in cell death.

Given these issues, OEHHA will continue to follow the recommendation by Judson et al. in that rather than using data flags as a hard filters for the relevancy of assay results, users can benefit from understanding that the flags for each assay may identify potential interference and possible sources of uncertainty. It is clear that flags should not be used to determine whether an assay is

active. Using flags and cytotoxicity as hard cutoffs to ignore assay results would exclude potentially important information.

The Appendix D in OEHHA's report summarizes the HTS assay results for the different dyes and has a column where the flags are all noted for each dye.
A section on the chemical quality of the different food dyes is now included in the revised draft report in Appendix A, which contains detail on the high throughput system assays. In terms of chemical analytical quality control, Blue No. 1, Red No. 3, and Yellow No. 5 have caution flags regarding their quality control. The results of assays on these dyes should not be excluded solely on caution flags; instead, the assay results should be assessed with a consideration of uncertainty. Although there may be numerous sample IDs for the same food dye color, the Chemistry dashboard only provided chemical QC data on food dyes that had a Tox21 ID. The following information on chemical quality is available from the Chemistry Dashboard:

Dye	Name	CASRN	DSSTox Substance ID	Tox21 ID	QC Grade
Blue No. 1	Brilliant Blue	3844-45-9	DTXSID2020189	Tox21_300516	T0; (Z) MW Confirmed, no purity info T4; (Z) MW Confirmed, no purity info
Blue No. 2	Indigo Carmine	860-22-0	DTXSID1020190	Tox21_113456	T0; (ND) Not Determined, Analytical analysis is in progress T4; (A) MW Confirmed, Purity > 90%
Blue No. 2	Indigo Carmine	860-22-0	DTXSID1020190	Tox21_302732	T0; (C) MW Confirmed, Purity 50-75% T4; (Fns) CAUTION, No Sample Detected; Biological Activity Unreliable
Green No. 3	Fast Green	2353-45-9	DTXSID3020673	Tox21_302086	T0; (A) MW Confirmed, Purity > 90% T4; (A) MW Confirmed, Purity > 90%
Red No. 3	Erythrosine	16423-68-0	DTXSID7021233	Tox21_202932	T0; (ND) Not Determined, Analytical analysis is in progress T4; (Cc) CAUTION, Low Concentration, Concentration 5-30% of expected value
Red No. 3	Erythrosine	16423-68-0	DTXSID7021233	Tox21_302085	T0; (ND) Not Determined, Analytical analysis is in progress T4; (F) CAUTION, Incorrect MW Biological Activity Unreliable

Dye	Name	CASRN	DSSTox Substance ID	Tox21 ID	QC Grade
Red No. 40	Allura Red	25956-17-6	DTXSID4024436	Tox21_300393	T0; (A) MW Confirmed, Purity > 90% T4; (A) MW Confirmed, Purity > 90%
Yellow No. 5	Tartrazine	1934-21-0	DTXSID1021455	Tox21_113411	T0; (F) CAUTION, Incorrect MW Biological Activity Unreliable T4; (I) ISOMERS, two or more isomers detected
Yellow No. 5	Tartrazine	1934-21-0	DTXSID1021455	Tox21_201539	T0; (A) MW Confirmed, Purity > 90% T4; (F) CAUTION, Incorrect MW Biological Activity Unreliable
Yellow No. 5	Tartrazine	1934-21-0	DTXSID1021455	Tox21_300554	T0; (A) MW Confirmed, Purity > 90% T4; (Ac) CAUTION, Low Concentration Concentration 5-30% of expected value
Yellow No. 6	Sunset Yellow	2783-94-0	DTXSID6021456	Tox21_201897	T0; (A) MW Confirmed, Purity > 90% T4; (A) MW Confirmed, Purity > 90%
Yellow No. 6	Sunset Yellow	2783-94-0	DTXSID6021456	Tox21_300407	T0; (A) MW Confirmed, Purity > 90% T4; (A) MW Confirmed, Purity > 90%

In reference to OEHHA's 2019 draft document on Acetaminophen, there is not an established standardized way of incorporating cytotoxicity or flags when interpreting ToxCast HTS data. As stated before, the ToxCast data pipeline is continuously evolving and changing, as new data is being incorporated. This could have resulted in changes to the parameters that contribute to defining hit calls.

Comment 11: (page 9) Animal Toxicology. "OEHHA reviewed a number of animal studies of potential relevance to the current assessment, concluding that "…*effects on activity, and learning and memory were reported in both young and adult animals*." However, OEHHA did not evaluate these animal studies for quality or reliability using any type of published method. In addition, for several studies the purity of the test substance was not reported, as such one cannot determine if the observed effect was due to the color additive or to an impurity. Much of this evidence has previously been evaluated (*e.g.*, by FDA, JECFA and EFSA). The lack of a robust, predictable animal model for assessing the effects of color additives on behavior has

been noted.[33]" The comment citation 33 is a "Quote from Dr. T. Scott Thurmond, Ph.D., FDA Review Toxicologist - "…*as far as I know the animal models are not the best choice for those types* [assessment of hyperactivity] *of studies*."; at an October 7, 2019 Science Board to the FDA meeting Expert Panel meeting, transcript available at https://www.fda.gov/media/135001/download "

Response 11: OEHHA's approach to evaluation of study quality includes the study characteristics assessed by published summary tools but does not depend entirely on those cursory methods.

Purity of test substances was provided in the majority (12 of 14) of the studies of individual dyes we reviewed. In other studies, the source of test substance was stated and purity could be found out by consulting the source. The FD&C batch certified colors are, of course, not 100% pure. The role of impurities in producing toxicity is always a topic to pursue once a toxic effect is identified for a test substance.

Regarding the citation [33], Dr. Thurmond was a toxicologist at FDA. He was speaking to the recommendations of the 2011 panel for more research in animal models. He said they did not follow this recommendation because:

> "Well, the FDA looked at the literature and decided that the animal was not a good model for assessing hyperactivity in humans or intolerance to certain compounds."

No background is given on the basis for this decision. Animal models of ADHD have long been in use and continue to be developed (Majdak 2016). More generally, activity is a major endpoint in USEPA neurotoxicity and developmental neurotoxicity screening guidelines. Animal models are always used with consideration of species differences, and in conjunction with *in vitro* and human data. However, they are as useful for toxicology determinations of hyperactivity as for other domains such as cancer, birth defects, and infertility.

This appreciation for the value of behavioral testing in animal models is reflected in the FDA Redbook (2000) **"**Behavioral testing has been established as a reliable toxicological index in safety assessment" "Behavioral testing can be readily incorporated into toxicity testing protocols and, together with neuropathological evaluation, can enhance the ability to assess neurotoxic hazard". "It should be clear that neurotoxic effects identified in experimental animal models may not always compare exactly with what may occur in humans. Nonetheless, these effects are still interpreted as being indicative of treatment related effects on the nervous system and predictive of possible adverse health effects in humans."

Responses to Comments from the Color Coalition

OEHHA received comments from the Color Coalition. The Coalition incorporated by reference the comment letter from the International Association of Color Manufacturers (IACM). (See our more detailed responses to comments submitted by IACM). The comments note that the Coalition and IACM worked together to review and respond to OEHHA's public review draft.

Comment 1: (page 1) "As a prefatory matter, we remind OEHHA that the U.S. Food and Drug Administration (FDA) has an extensive premarket approval and market surveillance program for the use of synthetic food colors. Any additional regulatory action by OEHHA or the California Legislature will create confusion in an area where the FDA has sole and preeminent responsibility. The potential patchwork of laws at the state level will generate confusion among

consumers. Clear, simple, and consistent national regulation informed by risk-based science will enhance consumer trust in these products. FDA currently provides this leadership."

Response 1: OEHHA is well aware of the FDA's approval process for food dyes. The objective of the report is to review the available scientific information on the potential for synthetic food dyes to affect children's behavior, and the report is non-regulatory in nature.

Comment 2: (page 2) "…robust reviews of the health impacts of synthetic food colors conducted by scientific bodies including the FDA and the European Food Safety Authority (EFSA) have generally found these ingredients to be safe for use as food additives. As such, the claims suggesting synthetic food colors cause possible attention deficit disorder / hyperactivity in children is not scientifically substantiated. Existing risk assessments by international bodies have dismissed and discounted much of the available neurobehavioral evidence in this respect. Rather, the basis and nature of the OEHHA risk assessment was precipitated by legislative interest and policy driven conclusions predicated on casual correlation."

"**Synthetic Food Colors are Recognized as Safe.** In March 2011, the FDA Food Advisory Committee (FAC), an expert panel of pediatricians, toxicologists, behavioral scientists, food scientists, and scientists in related fields, convened for a meeting to review all the available scientific data investigating a correlation between color additive intake and hyperactive behavior in children. After two days of scientific discussion, presentations by researchers, and public comment by parents and stakeholders, the FAC recommended that no warning label on products was needed to ensure the safe use of colors as food additives. The FAC concluded, based on all available evidence, that a causal relationship between the intake of synthetic color additives and hyperactivity in children could not be established."

"Specific to the colors OEHHA is reviewing, seven of the nine certified color additives (all but Orange B and Citrus 2 of which have little to no documented U.S. or international use) have been recently evaluated for their safety by international regulatory bodies such as the Joint FAO/WHO Expert Committee on Food Additives (JECFA) and by the EFSA, both of which have concluded that they continue to be safe for all ages, including children. Both JECFA and EFSA have evaluated and concluded that the available literature does not provide compelling evidence to raise any concern about impacts to ADHD or any neurobehavioral effects from consumption of synthetic colors."

Response 2:

OEHHA's findings are not predicated on Legislative interest, nor are they policy-driven. As regards other reviews of the information, OEHHA's review was more inclusive and rigorous. In 2011, US FDA asked its Food Advisory Committee (FAC) to consider available relevant data on the possible association between consumption of synthetic color additives in food and hyperactivity in children, and to advise FDA as to what action, if any, is warranted to ensure consumer safety. US FDA concluded that most children have no adverse effects when consuming foods containing color additives, but some evidence suggests that certain children may be sensitive to them.

The FDA states on their web page "*The FDA has reviewed and will continue to examine the effects of color additives on children's behavior. The totality of scientific evidence indicates that most children have no adverse effects when consuming foods containing color additives, but some evidence suggests that certain children may be sensitive to them…."*

As we describe in the Introduction, OEHHA did not limit the review to the question of effects on children diagnosed with ADHD or other behavioral disorders. Rather, OEHHA evaluated the epidemiological literature to determine whether there might be any effects on behavior of the FD&C batch-certified synthetic food dyes in children in the general population with or without a diagnosis of ADHD. We did not focus solely on effects related to activity and attention, but evaluated the literature for effects on other neurobehavioral impacts as well. In addition, OEHHA evaluated the animal toxicology literature relevant to neurological endpoints; these studies were not emphasized in the 2011 US FDA review (US FDA, 2011).

There is no documentation that is publicly available that the FDA (or JECFA or EFSA for that matter) reviewed the animal literature as thoroughly as did OEHHA. As can be seen in Tables 3.9 and 3.10 in the OEHHA report, there were a number of animal studies with statistically significant changes in behavior following administration of individual food dyes as well as mixtures of food dyes. Neurobehavioral effects were observed for Yellow No. 5, Yellow No. 6, Red No. 3, Red No. 40, and Blue No. 1. In the mixture studies, relevant neurobehavioral effects were also observed. The other authoritative bodies did not adequately weight these effects, in some cases because they viewed the database as a whole as being inadequate or not showing consistent effects. OEHHA agrees that more studies would be useful; however, the data cannot be simply dismissed as the other bodies appear to have done. There were statistically significant neurobehavioral effects, for some dyes in multiple studies. As regards consistency of effect, the studies used different protocols, dosing regimens, time to measuring effects and measured different domains of neurobehavior, thus precluding evaluation of consistency. Several well-conducted studies were published after the reviews of USFDA and in some cases after the updates from JECFA. Additional studies have been published since then, both in vivo and in vitro. Finally, we reviewed newer data relevant to mechanisms of action of potential neurobehavioral or neurotoxic effects of the food dyes. The authoritative bodies cited in the comment did not review and integrate all the mechanism, animal toxicology and human studies on the topic of neurotoxicity.

In short, OEHHA did a more thorough review of more scientific information than any of the international bodies noted in the comment. It should be noted that the EU requires that foods with azo dyes be labeled with a warning “may have an adverse effect on activity and attention in children”. (https://www.europarl.europa.eu/sides/getDoc.do?language=EN&type=IM-PRESS&reference=20080707IPR33563)

As far as safety, the ADIs are largely based on very old studies, published decades ago, and are not based on neurological outcomes. The studies that were the basis of the ADI, as noted in the comment, did not measure neurobehavioral effects. As well, the investigators only did very superficial screening for any neurotoxic effects. Many of the studies used by FDA to develop the ADIs are not publicly available in any detail to conduct a formal study quality review, as IACM contractors provide in Appendix A using the ToxR Tool for publicly available published studies. As such, the ADIs were based on old studies not published in any detail that lacked a thorough evaluation of neurobehavioral measures, did not measure outcomes that one would measure in a modern study assessing neurobehavioral effects, and could not be evaluated for study quality. Many of the studies cited by OEHHA found effects at the level of the NOAELs from these old studies used as a basis for the ADI, and in a few cases at the level of the ADI. IACM’s own contractors, while dismissive of many studies for what we consider inappropriate reasons (see comments below to specific studies), judged many of the studies OEHHA cites as “reliable without restrictions” or “reliable with restrictions” using the ToxR Tool and note a number of them scored high in their application of the ARRIVE guidelines (see Appendix A, Table A1).

Most importantly, there is substantial evidence from human studies that the food dyes impact behavior in some children. As noted above, FDA acknowledges this. While there was discussion that the neurobehavioral effects might be due to some “hypersensitivity”, that does not eliminate the concern, and no evidence that the effect is from an allergy type phenomenon is cited for this theory. OEHHA presents a thorough systematic review of the human literature in chapter 2 of the OEHHA report, including a comprehensive study quality evaluation. The clinical trial studies in children, where children were first given a diet free of food dyes and then provided food dye challenge, and a meta-analysis published in 2012 (Nigg et al, 2012) support an effect of food dyes on children’s behavior. One study (Stevenson et al., 2010) found that polymorphisms for histamine breakdown (histamine N-methyl transferase gene) are associated with children responding to dyes in the dye mixture studies. Interestingly, recent scientific studies have shown that histamine is a neurotransmitter and histaminergic neurons interact with many other structures in the brain including those involved in behavior (Yoshikawa et al, 2019; Scammel et al., 2019). OEHHA’s report states (in Section 2.8), after a thorough evaluation of the human clinical trials, “Based on the extent of the positive findings reported, and the fact that we could not convincingly or consistently attribute these positive findings to errors in study design or other bias, we conclude that the current human epidemiologic evidence supports a relationship between food dye exposure and adverse behavioral outcomes in children.”

Comment 3: (page 2-4) “**The Risk Assessment’s Underlying Assumptions are Misapplied.** Without establishing a neurobehavioral hazard for food colors, it is not possible to conduct a risk assessment for food colors, particularly one based on suggested neurobehavioral impacts in human trials burdened by significant limitations and confounders. OEHHA's draft report often emphasizes results from select *in vitro* studies that help support a presumed conclusion (i.e., that color additives affect behavior). At the same time, OEHHA diminishes other *in vitro* and *in vivo* data (e.g., Lok et al., 2013) that reach a different conclusion indicating a lack of evidence for neurobehavioral impacts. “

Response 3: Contrary to what is implied in the comment, OEHHA did identify a neurobehavioral hazard as presented in the assessment in chapter 5. The hazard identification is based on studies from multiple data streams, including clinical trials conducted in children, in vivo animal studies, in vitro mechanistic studies, with some albeit limited support from high throughput assay systems. Specific responses to comments critical of the human and animal studies are detailed below.

OEHHA did not have a "presumed conclusion”. Rather, we conducted a thorough review and analysis of all available information. It was after this thorough analysis that OEHHA drew conclusions based on the sum of the data reviewed. The *in vitro* data are discussed in the report, primarily in Chapter 4. While ToxStrategies identified studies not reviewed by OEHHA, they did not contain original or relevant findings (see comments and responses in section 3).

Comment 4: “In drawing its conclusions, OEHHA gives significant weight to non-guidance studies where weak statistical analysis is used to accentuate inconsistent signals. OEHHA also draws conclusions from "noise" in animal or in vitro studies and prioritizes such findings despite overwhelming evidence that supports a conclusion of no effect. Conversely, a lack of consistent results among studies generally leads to a weight-of evidence conclusion that an identifiable hazard does not exist.”

Response 4: OEHHA conducted a thorough review of the literature and a review of all animal studies obtained. We reviewed all of the clinical trials in children that met our inclusion criteria. We, in fact, did a very comprehensive review of the studies, including of the statistical analyses used and we note where the statistical analysis was flawed. We derive our conclusions from human and animal studies with sound statistical analyses, not "noise". The clinical studies in children, with appropriate and positive statistical findings support an effect of synthetic food dyes on behavior. We could not find a strong basis for dismissing these studies. As regards consistency of effect, the animal studies used different protocols, dosing regimens, time to measuring effects and measured different domains of neurobehavior, thus precluding evaluation of consistency.

The animal studies used as the basis of ADIs did not test for the types of effects observed in human studies, or any other specific neurological endpoints. Studies used by FDA and other bodies as the basis of their ADIs are many decades old and do not reflect the state of the art in either so-called "guidelines" neurotoxicology studies as we know them today or non-guideline neurotoxicology studies published in scientific journals. Note that because a study is a "guideline" study does not in and of itself mean it is a good study, nor that it measures outcomes appropriate to the question of whether food dyes impact neurobehavior. Contrary to the comment, there is not "overwhelming evidence that supports a conclusion of no effect".

Comment 5: "The majority of meta-analyses and systematic reviews of those meta-analyses published in the last 5-7 years have concluded that dietary intervention methods, including diet restriction approaches (including color restricting) and those that are pro-nutrient, do not significantly alter children's behavior. These conclusions do not support an association between food colors and neurobehavioral endpoints and should be appropriately considered within OEHHA's analysis and report."

Response 5: Please see our response to the comments from IACM regarding other systematic reviews and meta-analyses. Most of these dietary intervention studies are either not specifically focused on artificial food dyes, involve reviews of potential ADHD treatments, and/or do not include full evaluations of causal inference.

Comment 6: "The OEHHA report has a significant flaw regarding its inclusion of studies. While the report suggests that it has taken a systematic approach in its literature search and review, it does not describe the criteria used to qualify or exclude studies. This leaves the impression that OEHHA's weighting of studies in drawing conclusions is either arbitrary, selective for those that fit a narrative, or both."

Response 6: The comment that we did not look at the available information using a systematic review protocol is incorrect. OEHHA conducted a systematic review of the human studies, described in Chapter 1 and 2, and provided clear inclusion/exclusion criteria. The processes OEHHA used to perform its literature search are described in section 1.3. Detailed descriptions of the methods OEHHA used to qualify and exclude epidemiological studies are given in Sections 2.3 and 2.4. Because a number of broadly inclusive data sources were used to identify human studies, and because the same methods and procedures were applied regardless of study outcome, there is no reason to believe that these strategies led to a biased or arbitrary selection of studies. For animal studies, we conducted a thorough review of the literature and essentially retrieved and reviewed all articles, as there is not a large database for any of the

dyes. We used criteria similar to those in the ARRIVE guidelines (which focus on study reporting) for evaluating studies, and in our large tables of animal studies, we provide information that coincide with the areas included in the ARRIVE guidelines.

Comment 7: "Studies of mixtures of food colors are not appropriate for hazard identification. They do not allow the identification of specific food colors that might pose a hazard, if such a hazard exists. Additionally, many of the studies include color additives within the mixtures that are not approved for use in the United States nor within the scope of OEHHA's review. In fact, by considering combinations of colors, OEHHA has, in many cases, asserted effects for color additives that likely have no contribution to the identified hazard, if such a hazard exists at all within the study."

Response 7: OEHHA agrees that studies of mixtures present a problem with attribution of effects to any one dye. OEHHA states that, "Most studies involved administering multiple dyes at the same time so no single offending agent could be identified" (page 46). Regardless, these studies can still be used to evaluate whether the broad class of synthetic food dyes as a whole might contain at least some agents that cause adverse neurobehavioral effects. Moreover, children are not generally exposed to one dye at a time, but rather to multiple dyes simultaneously. As such, mixture studies reflect real world exposures. Indeed, almost all of the dyes were studied in humans as mixtures, with the exception of tartrazine (Yellow No. 5). The larger well-conducted mixture studies showed effects from the dye mixtures (see Table 2.1). In addition, there are animal toxicology studies of individual dyes that demonstrate effects on behavior and the results of the animal studies are part of our hazard identification.

Comment 8: "Clear evidence of causality must be present for risk management actions to be warranted and OEHHA does not provide any conclusive evidence.'

Response 8: OEHHA's report does not delve into risk management options. As regards causality, the epidemiologic literature we emphasize is comprised of human clinical trials that used a cross-over blinded design. These can be interpreted to provide strong evidence of causality by the nature of their design. We present a thorough evaluation of the human, animal, and in vitro studies available and synthesize the information. Together the available information supports that food dyes have neurobehavioral effects in some children. This conclusion is, in fact, not different from the FDA's conclusion that some children appear to be sensitive to food dyes. Also, our conclusion that the totality of the data supports that food dyes have neurobehavioral effects in children is not different than the EU's warning label, which requires that foods containing azo dyes be labeled with the warning "may have an adverse effect on activity and attention in children".

Comment 9: "In conclusion, it has been our repeated position that to single out synthetic colors as a focus of investigation is not a productive strategy for addressing an important disorder such as ADHD. The legislative interest and inquiry does not in and of itself establish significant questions of fact or dispute to suggest a hazard exists. The evidence, when appropriately contextualized, does not support discriminating against food colors within our national food safety program. We support federal regulations that result in uniform structures, empower consumers to make informed decisions, and are grounded in risk-based science. When a patchwork of regulatory policies exists, it contributes to consumer confusion and adds unnecessary stress to the supply chain resulting in higher prices."

Response 9: First, we did not single out synthetic food dyes to address ADHD. That was not the objective of the report. Rather, we did a very thorough analysis of all available data to

evaluate whether the synthetic food dyes impact neurobehavior in children. OEHHA's rigorous approach to evaluating all available information lead to the conclusion that the evidence supports an association between consumption of synthetic food dyes and neurobehavioral impacts in some children. As noted above, our findings were not influenced by Legislative interest.

Response to Comments from Center for Science in the Public Interest
The Center for Science in the Public Interest (CSPI) submitted a comment letter supplementing the letter submitted by 21 organizations and 31 researchers. In the letter, CSPI responded to criticisms of the OEHHA draft that were published on the International Association of Color Manufacturers (IACM) website regarding the OEHHA draft report.

Comment 1: (page 1)" IACM claims that "the assertions made in the draft report linking all FD&C colors with possible negative health or behavior effects are based on insufficient evidence."[1]

"This assertion flies in the face of the more than 200 relevant studies and reviews examined by OEHHA, including 27 clinical trials in children. Indeed, it is very rare for *any* human evidence to be available on the safety or toxicity of substances intentionally added to food, much less clinical trials, the "gold standard" for establishing causality. FDA does not require clinical trials or other human studies for food or color additives, and only recommends them "if indicated by available data or information."[2] FDA's guidance states, "However, petitioners may elect to perform such studies in certain circumstances, such as when the proposed additive will be consumed by humans at relatively high levels..."[3] In CSPI's experience, few elect to perform expensive clinical trials on substances added to food. The exceptions are for major nutrients such as fats or sugar or their substitutes (e.g., related to effects on weight, weight gain, or other metabolic parameters, or on gastrointestinal tolerance in the case of individual sugar alcohols or poorly absorbed sugars)."

"In addition to the 27 clinical trials meeting OEHHA's inclusion criteria, the majority of which found that children reacted behaviorally to synthetic food dyes, OEHHA also included numerous animal toxicology studies in its evaluation, as well as studies relevant to understanding the mechanisms by which synthetic food dyes might cause neurobehavioral effects."

"On top of the 230 studies related to human, animal, and mechanistic evidence it considered, OEHHA examined 283 assays in a sophisticated high-throughput assay evaluation that examined the extent to which dyes interact with specific molecular targets underlying neurological processes, or that otherwise trigger pathways suspected to lead to neurotoxicity."

"All three streams of evidence—human, animal, and mechanistic—support OEHHA's conclusion that "food dyes may cause or exacerbate neurobehavioral problems in some children." "

Response 1: The comment highlights a key element of our integrated approach to this assessment, the integration of evidence across human, animal, and mechanistic data. We believe that the preponderance of evidence across these studies, combined with our evaluation of exposure, supports our conclusions.

Comment 2: (page 2) "The evidence supporting OEHHA's conclusions that synthetic dyes cause or exacerbate neurobehavioral problems in children is consistent across evidence streams."

"IACM asserts that, "[t]he studies evaluated in the OEHHA risk assessment do not indicate consistent or strong associations between any food dyes and hyperactivity or neurologic effects in children."

"In fact, what is so striking about the OEHHA assessment—the first systematic review and integrated assessment ever conducted on synthetic food dyes—is how consistent findings are across evidence streams (human, animal, mechanistic). The findings in multiple evidence streams reinforce and strengthen OEHHA's conclusion that synthetic food dyes can cause or exacerbate neurobehavioral problems in children. As an example, we provide more detail on FD&C Yellow No. 5, below (Appendix 1)."

Response 2: See response above. In addition, there continue to be many important data gaps which should be addressed. There has been a limited amount of study of key outstanding questions related to the neurobehavioral impacts in recent years, particularly in the United States. We have highlighted some of these research needs in the document. Although more research would be useful, OEHHA's conclusion indicates the time has come to re-examine the widespread exposures to food dyes in children.

Comment 3: (page 2) "Positive associations between synthetic food dyes and neurobehavioral effects are consistently reported in children.

"The majority of studies in children actually do report evidence of an association between synthetic food dye exposure and neurobehavioral outcomes."

"Furthermore, as reported by OEHHA, studies using validated measures for assessing outcomes (generally higher-quality studies) were more likely to find positive associations than studies that did not use validated measures for assessing outcomes (70.6% vs. 50%), as were newer studies (conducted after 1990) compare to older studies (83.3% vs. 57.9%). This increases the confidence that the reported positive associations are real."

"In its handbook on conducting health assessments using systematic review and evidence integration, the National Toxicology Program considers "unexplained inconsistency" as a factor that can decrease confidence in the results. However, as indicated above, OEHHA's analysis reveals that inconsistencies seen between human studies can at least in part be explained by well known factors such as quality of the studies, size of the studies, and the size of the doses used."

The comment goes on to describe Bateman et al (2004) and McCann et al. (2007), which are well-conducted clinical trials studies using randomized double blinded cross-over design.

"…these studies addressed many limitations of earlier study designs, used several validated outcome measures for hyperactivity, and included both children with and without hyperactivity. In the first study there were significant ($P<0.001$) reductions in hyperactive behavior during the withdrawal phase of the study, and significantly greater ($p<0.02$) increases in hyperactive behavior during the challenge period compared to the placebo period. The effects were not influenced by the presence or absence of hyperactivity or of atopy."

"The second study validated the statistically significant effects on hyperactivity seen in the first study on three-year old children and extended the results to older children (8-9 years old). …Thus, the effects were replicated in three large samples of children (n=277, 140, 136)."

"No biases or other factors invalidate the positive associations reported. It is true that not every study conducted on food dyes for neurobehavioral effects found effects. This is not at all unusual and is generally the case when reviewing evidence on any topic. Studies differ in many

ways (e.g., number of subjects, age of subjects, sensitivity of subjects, dose of dyes, when exposure occurred, for how long, outcome measured, method used to measure the outcome), and all studies have limitations. OEHHA carefully examined these and other important features of the studies (e.g., adequacy of blinding, adequacy of placebo) to evaluate whether certain factors or biases might explain the results. …OEHHA concludes, "However, after extensive analyses we were unable to identify any clear set of biases or other factors that invalidated the positive associations reported in the current epidemiological literature.""

"… animal studies evaluated exposures in adult animals, animals exposed in-utero, and animals exposed from before birth through adulthood. In addition to differences in the timing and duration of exposure, different animal studies used different doses; different species; different behavioral tests with different sensitivities that measuring different behaviors in different ways; and measured effects at different times in the lifespan of the animal. Therefore, it should not be expected that results should be completely consistent."

"Importantly, studies that find no effects (a minority of the studies) do not outweigh studies that find effects….Furthermore, several meta-analyses have found that the pooled overall effect size is statistically significant in the direction of adverse effects."

Response 3: OEHHA evaluated the human clinical trials in a very thorough manner and we conclude that the studies support an association between food dye exposure and adverse neurobehavioral outcomes.

Comment 4: (page 4) "<u>OEHHA's conclusions are consistent with those of other independent reviewers.</u> OEHHA's assertions are fully in line with other independent reviews, including three metaanalyses, a review on behalf of the European ADHD Guidelines Group, a review using the Oxford Center for Evidence-Based Medicine guidelines, and others." (see original letter for citations)

Response 4: Thank you for pointing out consistencies with other independent reviews. OEHHA endeavors to operate independently of other organizations and to provide a thorough evaluation of the questions at hand.

Comment 5: (page 4) "<u>The association between food dyes and neurobehavioral effects is significant.</u> As noted previously, IACM asserts that, "The studies evaluated in the OEHHA risk assessment do not indicate consistent or strong associations between any food dyes and hyperactivity or neurologic effects in children."

"In fact, OEHHA evaluated the magnitude of the association using established criteria in each of the clinical trials in children that reported an association. Of the 16 out of 25 challenge studies that reported a statistically significant positive association, the magnitude could be measured in 13, and in 12 of those, the magnitude exceeded OEHHA's established criteria (the remaining one was borderline). In the remaining 3 studies, OEHHA found the magnitude of the statistically significant positive association to be unclear."

"The criteria used by OEHHA (20% or 0.20 effect size) is roughly the same as the effect size seen in the UK studies and obtained in the meta-analyses on dyes. While this effect size may

not be considered large from a statistical standpoint, the biological significance should not be underestimated."

"For example, subtle yet highly significant biological effects are seen in a large (N=11,640) study that followed children with behavioral problems of various causes for nine years using a reliable and well-validated measure of behavior; it found that each 1-point increase in inattention at age 7 was associated with a 6-7% increased risk of failing to gain five good grades on a standardized test taken by students aged 15-16 in England, Wales, Northern Ireland and other British territories."

Response 5: Your example illustrates an important point stressed in the OEHHA document, for an exposure which is widespread across a population even small effects noted in individuals may have large societal ramifications.

The commenters offer several recommendations.

Comment 6: (page 5) Convey Conclusions Clearly and Unambiguously; Avoid or Clarify use of "May" and "Some"

"The risk characterization section of the draft report (Chapter 7) states (p. 247),

> "Based on multiple streams of evidence, the FD&C synthetic food dyes cause or exacerbate neurobehavioral problems in children (see Chapter 5, Hazard Identification.)"

"CSPI fully supports this conclusion and recommends that this way of stating the conclusion be used throughout the report."

"We are concerned that some readers will misinterpret the conclusion in the Executive Summary, due to use of the term "may" and "some." That conclusion states,

> "The scientific literature indicates that synthetic food dyes may cause or exacerbate neurobehavioral problems in some children." We recommend that "may" and "some" be deleted."

"Saying "some" children implies that the neurobehavioral effects of synthetic food dyes are limited to a particular subpopulation of children. In fact, effects from synthetic dyes are not limited to children with food intolerances, ADHD, atopy, or any other identifiable subpopulation. OEHHA was not able to identify predictors for which children will experience behavioral problems from synthetic food dyes. The UK-funded studies were unable to identify any simple demographic or clinical attribute (e.g., gender, social class background, atopy) indicating which children would react. The UK team did identify certain genetic polymorphisms that appear to confer greater susceptibility to synthetic food dyes,[20] but further studies are needed to replicate the results, and it would not be practical to screen children for these genetic polymorphisms.

"Furthermore, the term "some" is redundant. There is always a distribution of response to exposures. Some people experience an adverse reaction to a drug, while others tolerate it well."

"… It is understood that the probability of developing an effect, disease, or health condition from an exposure will vary amongst individuals."

"This is important to clarify since it corrects an earlier view put forward by FDA at its Food Advisory Committee in 2011 that the effects of synthetic food dyes on behavior "appear to be due to a unique intolerance to these substances and not to any inherent neurotoxic properties."[21] In fact, the best animal studies and mechanistic evidence clarify that dyes do have neurotoxic properties."

""May" should be deleted since some readers interpret "may" to mean that OEHHA has not been able to arrive at a firm conclusion regarding the association between food dyes and neurobehavioral problems in children, because the evidence is not sufficiently strong. In our view, this incorrect. Moreover, it is not consistent with the conclusion in the risk characterization section of the report cited above ("the FD&C synthetic food dyes cause or exacerbate neurobehavioral problems in children"). Because results are consistent across evidence streams, and include human evidence derived from 27 clinical trials, OEHHA has reached a firm conclusion, based on sufficient evidence, and this should be made clear."

"In our view, the use of "may" in this context does not reflect doubt or uncertainty about whether or not there is an association between synthetic food dyes and neurobehavioral problems in children, but reflects the fact that the response of people varies and that not each and every person will experience neurobehavioral problems when exposed to synthetic food dyes. This is analogous to use of "may" on drug labels, for example, some antihistamines that state, "may increase drowsiness," not because there is doubt about whether antihistamines cause or increase drowsiness, but because responses between people may vary (and responses may vary even for the same person, depending on other factors)."

'As stated above, we recommend that OEHHA use the form of the conclusion in its risk characterization, i.e., "Based on multiple streams of evidence, the FD&C synthetic food dyes cause or exacerbate neurobehavioral problems in children." Alternatively, the word "may" in the Executive Summary conclusion could be replaced by "can." If OEHHA uses the term "may," it should clarify that "may" does not reflect the level of confidence in the results or the sufficiency of the evidence, but rather, is an acknowledgement that responses to dyes, just as to any type of exposure, from smoking to COVID-19 to medications, can vary. And, for the reasons stated above, we recommend that the word "some" before "children" be deleted."

Response 6: The point that the exact wording of statements from OEHHA may have significant impacts on how conclusions are interpreted is well taken. "May" has been used in the sense that exposures to food dyes can cause measurable neurobehavioral impact, but, as the comment points out, not all children's reaction are the same. This was not meant to detract from the firmness of our conclusion. We agree that "can" may be more appropriately interpreted by the interested public as consistent with our intent, and we will make that wording change.

We think that the use of the word "some" to indicate that effects are identified in "some children" is the appropriate way to describe the finding. While the commenter has a clear understanding that there are often susceptible and unsusceptible populations to most effects, this may not be the case for many readers. In studies of food dyes, only a minority of children are identified in epidemiologic studies as responding to the dyes. The comment is correct that we do not have an ability as yet to easily clinically identify these children. For the purpose of accuracy, we

believe that including “some” as a modifier for children in this context is appropriate for this document.

Comment 7: (page 7) “Consider the cumulative effect of color additives, taking into account chemically or pharmacologically related substances, though not at the expense of delaying the final report.”

“Studies of synthetic food dyes have been conducted both on individual dyes and mixtures of dyes. For example, FD&C Yellow No. 5 is the only dye that has been tested individually in humans. Studies of mixtures of synthetic food dyes have been conducted with the recognition that children are not exposed to dyes in isolation but rather are exposed to mixtures of synthetic food dyes from multiple sources, primarily food, drugs, and supplements, that may have interactions (e.g., synergistic or additive effects).”

“FD&C Yellow No. 5, Yellow No. 6, and Red No. 40 are all chemically related azo dyes. These three dyes comprise more than 90% of dyes certified for use in food in the United States. In many studies of mixtures of dyes in children, the amount of dye used was proportionate to the amount certified. Thus, these three azo dyes dominated many of the mixtures tested.”

“Section 706 of the Federal Food, Drug and Cosmetic Act codified under 21 U.S.C. § 379e states:

> “In determining, for the purposes of this section, whether a proposed use of a color additive is safe, the Secretary shall consider, among other relevant factors—
>
> (i) …
>
> (ii) the cumulative effect, if any, of such additive in the diet of man or animals, taking into account the same or any chemically or pharmacologically related substance or substances in such diet; …”

“Furthermore, 21 CFR 70.11 states that in the absence of evidence to the contrary, color additives that cause similar or related effects should be considered to have additive toxic effects:

> Different color additives may cause similar or related pharmacological or biological effects, and in the absence of evidence to the contrary, those that do so will be considered to have additive toxic effects.”

“In characterizing the risks of dyes, OEHHA should develop appropriate reference levels based on neurobehavioral effects, and consider the cumulative effect of each dye, taking into account chemically or pharmacologically related substances, especially including other synthetic food dyes. At the very least, the cumulative effect of chemically related azo dyes should be considered. A number of resources may be useful in this regard.”

Response 7: The impact of cumulative effects of exposures on children is certainly a topic of interest at OEHHA. It is beyond the scope of our project to evaluate the cumulative effects of food dyes on children. The project also did not extend to other potential impacts of these chemicals.

The U.S. FDA has regulatory purview over the FD&C batch-certified food dyes that are the subject of our evaluation. We are not developing reference levels or ADIs for these dyes. We do point out in our document a need to revisit the ADIs, which would be the responsibility of the FDA.

Comment 8: "Given that there are no human trials on FD&C Red 40 and Yellow 6 individually, should there be any doubt as to whether those dyes cause or exacerbate neurobehavioral problems in children—despite the ample animal and mechanistic evidence on those dyes when tested individually, and the evidence from human and animal studies of the neurobehavioral effects of mixtures in which those dyes (plus FD&C Yellow No. 5) contributed the largest amounts to the mixtures—then the compelling evidence on FD&C Yellow No. 5 demonstrating neurobehavioral effects could be utilized for FD&C Red No. 40 and Yellow No. 6, since they are in the same chemical class (azo dyes). This is analogous to the decision FDA made to apply data demonstrating reproductive and developmental toxicity on PFOA to other long chain perfluorinated chemicals for which such data were unavailable."

Response 8: Management of chemicals by class, when appropriate as in your example, is certainly an approach of interest at OEHHA. The azo dyes are a candidate for consideration of this approach for regulatory management. That question, however, is beyond the scope of our current assessment. As noted in our response to comment 7, the regulation of food dyes is in responsibility of the FDA.

Response to Comments from 21 Organizations and 31 Researchers

OEHHA received comments from a group of organizations and researchers largely supportive of the assessment.

Comment 1: (page1) "OEHHA's draft report was informed by expert scientists attending a public scientific symposium organized by OEHHA, and by the public who were invited to submit information. The process was and is transparent, inclusive and resulted in a well-written draft report with clear, robust conclusions.

"Overall, the draft assessment concludes that:

"The scientific literature indicates that food dyes may cause or exacerbate neurobehavioral problems in some children."

"We therefore strongly agree, as the draft assessment notes, that:

"At a minimum, in the short-term, the neurobehavioral effects of synthetic food dyes in children should be acknowledged and steps taken to reduce exposure to these dyes in children."

"The OEHHA draft assessment breaks new ground in **seven important ways:**

"First: The draft health effects assessment uses a state-of-the-art structured approach of systematic review and evidence integration that has never been undertaken on the neurobehavioral effects of synthetic food dyes. Systematic review and evidence integration methodology and structured frameworks are increasingly recommended by a wide range of agencies and institutions to address environmental health questions, to ensure collection of the most complete and reliable evidence, allowing them to be integrated to form the basis for conclusions."

"OEHHA's approach involved a rigorous, well-designed literature search strategy, systematic evaluation of study methods and quality to ensure an emphasis on studies of high quality, a thorough and exacting evaluation of the human, animal, and mechanistic evidence, and the development of conclusions both within and across evidence streams. Confidence in the analysis is supported by the fact that three evidence streams (human, animal, mechanistic) support the important conclusions reached: that synthetic food dyes can cause or exacerbate neurobehavioral effects in children."

Response 1: We have attempted to include and integrate all of the key evidence generated by multiple disciplines. Research has been conducted using a large variety of approaches and outcomes studied over multiple decades of research. We believe that the preponderance of evidence across these studies, combined with our evaluation of exposure, supports our conclusions.

Comment 2: (page 2) "Second: OEHHA's evaluation of human evidence is very stringent; we note that the amount and quality of human evidence on synthetic food dyes is arguably among the greatest of the thousands of substances intentionally added to food. It is worth underscoring that for the human evidence, the OEHHA assessment relies exclusively on controlled clinical trials, considered the "gold standard" of study design for evaluating causality. Rarely, if ever, are clinical trials available for additives or other substances added to food. In fact, 27 clinical trials in children met OEHHA's inclusion criteria. We thus agree that "high confidence" is warranted for conclusions from the results of these human studies, and that

there is “a substantial amount of evidence that consumption of synthetic food dyes is associated with adverse neurobehavioral outcomes in children.””

“Our confidence in the conclusions is further bolstered by the fact that OEHHA was unable to identify, after extensive analyses, any clear set of biases or other factors, such as inadequate blinding or randomization, that may have invalidated positive associations reported in the epidemiological literature.”

“Since excellent meta-analyses of the human evidence on behavioral effects of synthetic food dyes had been published recently, OEHHA sensibly did not perform a full metaanalysis, and instead undertook an unprecedented evaluation to examine strengths and weaknesses and trends in the evidence, considering many important factors including blinding, dose levels, and wash-out periods, as they may have compromised study results. OEHHA conducted its own statistical analyses to verify or add to reported associations and effect sizes. We fully support this approach.”

Response 2: We agree that for most chemicals that OEHHA examines, data from clinical dosing trials are not available. The information garnered from these studies, when well conducted, is exceptional in the world of environmental health research and risk assessment and provides strong evidence of causality. OEHHA evaluated the human clinical trials in a very thorough manner and we conclude that the studies support an association between food dye exposure and adverse neurobehavioral outcomes.

Comment 3: (page 2) “Third: Prior to this review, no such structured and robust systematic reviews of the animal and mechanistic evidence have ever been conducted on the neurobehavioral effects of synthetic food dyes. The assessment contracted by the U.S. Food and Drug Administration (FDA) for its Food Advisory Committee meeting in 2011 included only a two-page overview of animal studies, and the animal evidence was not discussed by the committee.”

“OEHHA’s assessment of animal studies was far more extensive and rigorous. In addition, it includes studies published since 2011 not available to FDA that are relevant to an assessment of neurobehavioral effects. OEHHA’s assessments of the animal evidence on specific food dyes, as well as on mixtures of dyes, support the conclusions from the human evidence that dyes can cause neurobehavioral effects in children. They also bridge some of the gaps in understanding from the human evidence, which is largely focused on the effects of mixtures of food dyes resulting from exposures in children. Overall, the animal literature indicates that synthetic dyes can affect activity, memory and learning, and effect changes in neurotransmitter systems in the brain and microscopic changes in brain structure in animals.”

“Similarly, FDA’s contractor included only a two-page discussion of possible biological processes/mechanisms. In contrast, OEHHA undertook an extensive evaluation of dyes tested in high-throughput assay systems, finding that there is evidence for a number of ways that neurobehavioral effects of synthetic dyes might occur, including interaction of food dyes with neurotransmitter systems.”

“The only other such review of the mechanistic data was funded by the American Beverage Association and was far less comprehensive, using fewer assays and using data on cytotoxicity to exclude potentially important information (i.e., considering certain results to be false positives that might not be) rather than flagging it for further consideration.”

Response 3: OEHHA recognized that an integrated approach was essential for an evaluation of potential risks. For the assessment:

- OEHHA evaluated the literature on human studies relevant to whether behavior is affected in children when they consumed synthetic food dyes.
- OEHHA evaluated the literature relevant to neurobehavioral effects in laboratory animals following synthetic food dye exposure.
- OEHHA evaluated information relevant to how synthetic food dyes might exert neurobehavioral effects, including data obtained through high-throughput screening assays (laboratory tests that evaluate the effects of chemicals on cells or biological molecules) conducted by the US Environmental Protection Agency and its federal agency partners.
- Along with collaborators at the University of California's Berkeley and Davis campuses, OEHHA estimated exposures to each FD&C batch-certified synthetic food dye in general use in the US for children of varying age groups as well as for pregnant women and women of childbearing age.
- OEHHA conducted a risk characterization where we present a number of comparisons to gauge whether exposure to food dyes may present risk of neurobehavioral impacts.

At the onset of our review process OEHHA conducted an open data call-in. In September 2019, OEHHA conducted a public one-and-a-half day symposium that was webcast (and is still available for archived viewing) on the neurobehavioral impact of synthetic food dyes. We assembled experts on exposure, toxicology, FDA regulatory history, epidemiology, and clinical medicine to review the science. This informed our process in drafting the document. Our draft document has been circulated to the public and we have received public comments to which we are responding. In addition, we initiated a peer review managed by the University of California. The peer review comments are public and our responses to the peer review comments are part of our public response to comments. Our process been rigorous, open, and transparent.

Comment 4: (page 3) "Fourth: The report includes the most up-to-date and comprehensive exposure assessment ever conducted on synthetic food dyes in the United States, building on that conducted by the FDA. The OEHHA analysis is the only assessment to evaluate U.S. exposures to synthetic food dyes in pregnant women, women of child-bearing age, and in as many different age groups of children (0-<2 years, 2-<5 years, 5-<9 years, 9<16 years, 16-18 years). This is noteworthy since these populations are suspected of being at greater risk from synthetic food dyes, compared to the general population. "

"In comparison with FDA's exposure assessment,[3] OEHHA's assessment also looked at more population/age groups, used more recent NHANES food consumption data, covered a broader distribution of estimates (mean, median, 75th and 95th percentiles), and incorporated food dye concentration measurements from University of California, Davis (UC-Davis) as well as from the FDA. Because it was assessing neurobehavioral effects that result from short-term exposures, as documented by numerous clinical trials, it was appropriate that the assessment used single day and 2-day averages of food consumption, rather than the 2-day and 10-14-day data previously used by FDA."

"We took note that the highest 95th percentile single-day exposure estimate was for FD&C Red No. 3 in children under 2, an age group that FDA did not evaluate in its exposure assessment."

"OEHHA's analysis found that beverages were the dominant sources of exposure for children to the dyes that contributed most to exposure to synthetic food dyes: FD&C Red No. 40, Yellow No. 5, and Yellow No. 6. Specifically, fruit juice drinks and soft drinks were the primary sources

of exposure for FD&C Red No. 40 and Yellow No. 6, and powdered fruit flavored drinks and fruit juice drinks for FD&C Yellow No. 5 in children."

"In addition, white foods, brown foods, and other foods that many do not associate with food dyes were found to be major contributors to food dye intake, including chocolate pudding and fruit muffins— important sources of FD&C Blue No. 2 for children under 2—and white icing, the largest contributor to Blue No. 1 exposure for children under 5, and a primary source of exposure to FD&C Red No. 3. These findings illustrate some of the challenges that consumers face in identifying foods containing synthetic dyes, as consumers are unlikely to expect that such foods contain dyes."

"Beverages, including juice drinks, fruit-flavored drinks from powders that get reconstituted, and soft drinks should be a priority for exposure reduction efforts, since they are among the products most commonly associated with food dye exposure."

Response 4: Within our limitations of funding and scope of work, OEHHA did confirm previous exposure assessments and update national exposure assessments as noted. Certain foods stand out as prominent exposure sources and these vary by age groupings. Since the objective of our assessment is to evaluate the potential for adverse neurobehavioral impacts in children, it was important to evaluate exposures in various age groups of children, and for women of childbearing age and pregnant women to gauge consumption of food dyes during in utero development.

Comment 5: (page 4) "Fifth: The report provides the first substantial data on socioeconomic differences, such as poverty level and race, that have long been suspected to impact levels of exposure to synthetic dyes. It is concerning that Non-Hispanic Black children (0-18) and women of childbearing age (18-49) had significantly higher intake of synthetic food dyes compared to their non-Hispanic White or Asian counterparts; and total food dye exposures were significantly higher among women of childbearing age with lower income, when compared to women with higher income."

"Lower income consumers may face greater challenges in avoiding synthetic food dyes than do higher income consumers. They are more likely to lack or have limited access to grocery stores or other stores that carry the same range of products available in more affluent communities. In addition, more affluent neighborhoods are more likely to include supermarket chains such as Whole Foods or Trader Joes that have explicit policies barring dyed foods, or others that carry "free-from" brands that prohibit food dyes."

"This is just one reason why putting the burden on consumers to avoid synthetic food dyes is unlikely to be fully effective and raises real equity issues. It is also important in that lower income communities experience higher rates of neurodevelopmental disorders like ADHD, as well as a concentration of other risk factors.[7,8] The accumulation of such risk factors likely contributes to the increased incidence of diagnosable ADHD in these socially disadvantaged groups. Therefore, to better protect these and indeed all populations at risk, risk factors that can be addressed, should be, to reduce cumulative population risk and cost and aggravating social/racial inequities. The fact that synthetic food dyes are a risk factor whose only function is cosmetic makes it all the more compelling."

Response 5: OEHHA did not identify other literature specifically addressing consumption variability by socioeconomic factors, and decided to include it in our analysis. Additional study

by others of differential exposure by race and ethnicity, economic status, and other factors is suggested as a research priority to better understand this issue.

Comment 6: (page 4) "Sixth: The assessment is the first to conduct a risk characterization comparing the basis for current "safe" levels established by the FDA between 35 and 70 years ago, called acceptable daily intakes (ADIs), with the results from studies relevant to the neurobehavioral effects of synthetic food dyes. We agree with OEHHA's conclusion that: "[t]he animal studies that form the basis of the FDA ADIs are many decades old and were not capable of detecting the types of neurobehavioral outcomes measured in later studies, or for which there is concern in children consuming synthetic food dyes." This critical fact was pointed out as well by members of FDA's Food Advisory Committee back in 2011."

"The OEHHA analysis makes a further contribution by providing quantitative estimates showing how far off the mark the FDA's current "safe" levels are. Remarkably, according to the report, the ADIs that FDA set for FD&C Red No. 3, Red No. 40, Yellow No. 5, and Yellow No. 6, are 10 to 100 times higher than they would be if they were based on the results of more recent studies that are appropriate for evaluating neurobehavioral effects. OEHHA notes that in the case of FD&C Red No. 40 and Yellow No. 5, they could be 1,000 times lower, depending on the method used to estimate the ADI."

Response 6: In Section 7.3, we compared the effective doses in reviewed animal studies to the NOAELs used by US FDA to derive human ADIs (see Tables 7.2a and 7.2b). As well, in Section 7.5, we compare NOAELs from several of the reviewed studies that could be used as a basis for setting an ADI for neurobehavioral endpoints to the existing FDA ADIs. These comparisons make the point that current ADIs are not adequately protective of neurobehavioral impacts in children.

Comment 7: (page 5) "Seventh: OEHHA's efforts also resulted in the first publicly available data on the amounts of food dyes in over-the-counter (OTC) medications and supplements marketed to children and pregnant women in the United States. FDA never has meaningfully considered the contribution of non-food sources of food dyes to overall exposure, although federal law requires the agency to consider "the probable consumption of, and/or other relevant exposure from the additive and of any substances formed in or on food, drugs, or cosmetics because of such additive; and the cumulative effect, if any, of such additive in the diet of man or animals…."[10]

"Researchers from UC-Davis tested national brands of OTC medications and supplements sold in well-known retailers (Target, CVS, WalMart, and RiteAid) in three states, testing samples from three different lots for each brand. The results reveal, as noted in the report, that some brands, and in particular some grape-flavored pain-relief syrups and cough/cold/allergy syrups marketed to children, can contain high levels of dyes, particularly FD&C Red 40, but also Blue 1. Following the label instructions can result in intakes that exceed FDA's outdated ADIs (even without accounting for any intake of synthetic dyes from foods)."

"OEHHA found that OTC medications can produce 10 to 40 times higher exposures to FD&C Red 40 on a given day than does typical consumption from food. Although consumption of most OTC medications may be infrequent, some (e.g., those for allergies) may be used more frequently, almost daily in certain seasons. Even for OTC medications given infrequently to sick children, it is unacceptable to include (for cosmetic purposes only) high amounts of agents that may cause or exacerbate neurobehavioral symptoms. "

Response 7: Our evaluation of the possible contribution to children's exposure to food dyes did include a first ever analysis of food dye content in certain children's OTC pharmaceuticals. We draw your attention to findings that have been changed in the final version of the document. We were contacted by Consumer Healthcare Products Association (CHPA), who found an error in the publication we used to estimate exposure to synthetic food dyes through OTC cough/cold/allergy and pain relieve/fever reducer syrups for children. The error is limited to the syrup formulations and did not affect estimates from other formulations. In response to the comments from CHPA, we confirmed the inaccuracies in our initial color additive exposure estimates for children from cough/cold/allergy and pain relieve/fever reducer syrups, and have made changes to the text and tables of the report. With one exception, the corrected values are all 100-times lower than initially reported and **none of the estimated exposure values for OTC syrups exceed the US FDA or JEFCA ADIs.** We believe we have addressed each of CHPA's comments and concerns. Please see the response to comments from the CHPA for details of the changes in the document.

Comment 8: (page 5) "It is not possible to predict which children will react negatively to dyes. OEHHA considered multiple factors that might influence susceptibility to synthetic food dyes, but concluded that "while it seems likely that sensitive populations exist, we did not find evidence that there is a simple and accurate way to identify these particularly sensitive children." For example, OEHHA's assessment considered children in the general population, with or without a diagnosis of ADHD or hyperactivity, and found that studies including only such children with such a diagnosis were not more likely to report an association with adverse impacts of food dyes."

"For this and other reasons, efforts to reduce exposure to synthetic food dyes should be directed at the population at large. Indeed, the entire concept of isolating a susceptible group is suspect. All environmental exposures affect people variably, whether the exposure be smoking, lead exposure, or COVID-19. That some people are more susceptible than others should not be a consideration in determining whether protective measures are needed for the population. Only population-wide measures can protect the community's sensitive members at this point."

Response 8: It is currently not possible to identify which children will respond to the food dyes. Stevenson et al. (2010) reports that polymorphisms in the histamine metabolizing enzyme histamine N-methyl transferase was associated with children's response to food dyes. Other than this association, OEHHA was unable to identify studies that found factors making children more susceptible to food dyes. As well the clinical trial studies in children indicate it is not just children with identifiable problems such as ADHD who respond to the synthetic food dyes – some of the children from the general population who responded to the food dyes had not been diagnosed with identifiable behavioral problems.

Comment 9: (page 6) "The effects of food dyes go beyond short-term effects on activity and attention and can have long-term, serious consequences. In addition to hyperactivity, inattentiveness, and restlessness, studies in children report other neurobehavioral effects of synthetic food dyes, including sleeplessness, irritability, and aggression.

While the effects of dyes may be short-term, and resolve after discontinuation of exposure, given the prevalence of synthetic food dyes in foods, supplements, and drugs, it is likely that exposures and, therefore, the related effects will occur repeatedly, impacting the ability of children to learn, succeed at school, and get along with peers on an on-going basis, with serious long-term consequences. The symptoms of synthetic food dye exposure overlap with ADHD-

type symptoms, and ADHD itself is associated with lifelong impairment in functioning and long-term outcomes that can include academic achievement (e.g., failure to complete high school), serious substance abuse, criminality, and depression."

"Evidence on effects from in utero exposure in animals raises concerns about pregnant women's exposure to synthetic dyes. While clinical trials have focused on adverse behavioral impacts from short-term exposure to synthetic food dyes in children, animal studies indicate that food dyes can also cause effects from in utero exposure, including for levels of food dyes found to have no effects in the (outdated and insensitive) studies used by FDA."

"Furthermore, these effects can be long-term, including effects on activity in the animals as adults. As the report notes, "the finding of behavioral and tissue marker effects of in utero- only exposure detected long after discontinuation of treatment speaks to an interference with developmental processes. More research would be needed to define a mechanism pathway from the tissue assays."

"Animal studies provide some evidence of impacts due to adult exposures. Animal studies find neurobehavioral effects resulting from exposures in adults at levels much lower than those reported to cause toxic effects in studies used as the basis of FDA's ADIs. These results suggest that an approach limiting exposure to synthetic food dyes to populations *in addition* to children may be prudent."

Response 9: In terms of long-term effects, we were not able to identify any human studies examining potential long term effects of exposure to synthetic food dyes. Some animal studies cited in the document identify outcomes which suggest the possibility of long term impacts from early life exposures. We have added text to the summary section 8.6 in Chapter 8 about the need for research into long-term effects of repeated exposure to synthetic food dyes.

Comment 10: (page 6) "There is no convincing rationale to exempt any of the dyes examined from efforts to reduce exposure. Much of the focus on synthetic dyes and neurobehavioral effects has centered around azo dyes, a chemical class of synthetic dyes which includes FD&C Red 40, Yellow 5, and Yellow 6, which together comprise over 90% of the dyes certified for use in foods in the U.S."

"However, we note OEHHA's conclusion that FD&C Red No. 3 has at least as much evidence as FD&C Red No. 40 and Yellow 5 (these three have been the subject of more studies)."

"In addition, according to the draft report, "Red 3 and Green 3 had hits for all the neurorelevant molecular targets that they were tested in" in the high throughput screening evaluation. Similarly, the OEHHA report finds that "[a]ll the FD&C synthetic food dyes (except for yellow dyes) are active for antagonistic effects with the thyroid hormone," which, it explains, may be linked to developmental neurotoxicity. It also found that all dyes were active in assays targeting dopaminergic and opioid receptors."

"OEHHA acknowledges that these are several pertinent associations between dyes and certain molecular targets of interest."

"In sum, there does not seem to be a sound basis for concluding that any of the seven dyes examined should be singled out as lacking evidence of neurobehavioral effects."

Response 10: We agree that there is not a sound scientific basis for concluding that any of the dyes we examined should be deemed to lack neurobehavioral effects. The evidence of effects in children in studies of dye mixtures makes it difficult to understand which of the dyes are the

worst actors. Individual dyes were studied in animal models and there is some evidence of effects on the nervous system of each of the dyes individually (with the possible exception of Blue No. 2, which has not been individually studied), including many measured effects on behavior, in these animal studies. OEHHA's assessment reviewed animal studies, both in vivo and in vitro, and high throughput assays that identify a number of potential pathways by which various synthetic food dyes may exert a neurological effect. While pointing to potential pathways by which a dye may exert a neurobehavioral effect, the mechanistic information is not robust enough to clearly identify mechanisms for all of the dyes or even any individual dye.

The literature we reviewed was pursued by independent investigators with specific interests and sources of support. FDA does not require a submission of a consistent database of pre-certification studies that includes neurotoxicity or an evaluation of neurobehavioral effects and we did not identify any long term basic research funding here in the US to address this issue. In fact, most of the recent research we reviewed has been conducted internationally. As described by the commenter, this database does not exclude any one dye as a potential actor in the dye mixture studies in children. At the same time, it does not identify any dye as the only or primary actor. Most certainly this literature does not lead to the conclusion that food dyes are not neurobehavioral toxicants. The section on research gaps in the OEHHA document helps outline in the information needed to further elucidate the toxicity of individual dyes and their impact individually and as a group on children's behavior.

Comment 11: (page 7) We urge that OEHHA expand on its conclusion that steps be taken to reduce exposure to these dyes in children by investigating effective ways to reduce exposure to synthetic food dyes, both in foods and in OTC medications and supplements. This is needed to ensure that its important findings are translated into measures that effectively reduce exposures and provide meaningful public health protection for children.

Response 11: Risk management suggestions are outside the scope of the report. OEHHA evaluated the scientific information relevant to whether the synthetic food dyes impact neurobehavioral outcomes in children.

Comment 12: (page 8) "We recommend that OEHHA further clarify that efforts to reduce synthetic food dyes should not wait for additional research. We wish to emphasize, as OEHHA states, that additional research is "a long-term proposition," and add that at this juncture, waiting, perhaps indefinitely, for even more evidence is inappropriate. Additional research is always generally desirable, including in the five areas that OEHHA describes."

"But a laudable desire for better data does not justify indefinite delay in using the data we have available. It is clear that there is more than sufficient evidence, from the 27 clinical trials as well as the animal and mechanistic evidence streams, to conclude that synthetic food dyes can affect children, as OEHHA has concluded. Efforts should be made "in the short-term," to reduce exposure to synthetic food dyes, as recommended by OEHHA. Indeed, the burden of proof should be shifted—food dyes should be reduced in the population until further research can provide evidence of their safety sufficiently strong to over-ride the extensive evidence assembled to-date."

Response 12: In the conclusions (section 8.6), OEHHA makes recommendations for areas of research that would help elucidate the effects of food dyes on children. However, there is enough evidence now that the synthetic food dyes impact children, as we state in our conclusions. We have clarified this in Section 8.6 where we make research recommendations.

It is not in the scope of the project to make specific recommendations on potential actions to reduce exposure.

Comment 13: (page 8) "We recommend that OEHHA include more information on exposure to mixtures of dyes, though not at the expense of delaying the current final report. Children and other consumers of synthetic food dyes do not consume food dyes one at a time; they consume mixtures of synthetic food dyes, over the course of a day, over the course of a meal, and even in a single product. Furthermore, clinical trials typically test mixtures of food dyes. However, OEHHA focuses mostly on exposure estimates of individual dyes. The information that is presented on total dye exposure is presented for children over a large range of ages, which has limitations since children vary so much in body weight over different ages. It would be helpful to know how the exposures of children of different age groupings to mixtures of dyes compares to clinical trials testing mixtures of dyes for certain age groups."

Response 13: Although it would be interesting to see this analysis, the data that OEHHA could access would not be adequate to conduct this analysis.

Comment 14: (page 9) We recommend that OEHHA characterize risk by comparing estimated exposures to short-term reference levels based on appropriate studies relevant for assessing neurobehavioral effects of food dyes. Currently, OEHHA compares estimated exposures to Acceptable Daily Intakes established by FDA and the Joint Expert Committee on Food Additives (JECFA). However, as OEHHA acknowledges, these do not characterize the risk of neurobehavioral changes. OEHHA should therefore develop more appropriate reference levels to use in characterizing risk.

Response 14: In the document, OEHHA examined the potential for change in current Acceptable Daily Intakes (ADIs) if modern studies examining neurobehavioral outcomes were utilized. None of the current ADIs are based on these outcomes. OEHHA does not have within its mandate the creation of reference levels for food additives. Rather, the US FDA is mandated to establish ADIs.

Response to Comments from Environmental Working Group

Comment 1: "EWG supports OEHHA's timely review of synthetic food dyes and agrees with OEHHA's conclusion that they "may cause or exacerbate neurobehavioral problems in some children." "

"… Our organization has long advised parents and families to limit children's consumption of these food additives, because of their potential to cause behavioral effects and other health harms. The current legally allowable Acceptable Daily Intake, or ADI, for food colorants was set by the Food and Drug Administration decades ago, based on toxicological studies that are now 35 to 70 years old. These earlier studies in laboratory animals were not capable of detecting the types of neurobehavioral identified by contemporary epidemiological research. Thus, we wholeheartedly agree with OEHHA's conclusion that the legally allowable ADIs do not protect children's health."

"In its health assessment of seven synthetic food dyes, OEHHA found that these dyes are associated with adverse neurobehavioral outcomes in children and that children vary in their sensitivity to exposure. These findings were supported by data from epidemiological studies, neurotoxicity studies in animals, and high-throughput and *in vitro* assays. Studies in humans showed that dyes were associated with the neurobehavioral effects of inattentiveness, hyperactivity and restlessness in sensitive children. In studies in adult animals, exposure to Red No. 3, Red No. 40, Yellow No. 5 and Yellow No. 6 resulted in altered brain chemistry. to Red No. 40 and Yellow No. 5 was associated with altered learning and memory and changes in brain cells, and exposure to Red No. 3 and Yellow No. 5. resulted in changes in activity and brain chemistry."

"From OEHHA's exposure assessment, the highest exposure was from Red. No. 40, followed by Yellow No. 5 and Yellow No. 6. It was found that dyes in children's over-the-counter medicine and vitamins were a source of higher exposure compared to food, especially for Red No. 40. Several doses of medicine in a day could lead to exceeding the current legally allowable daily intake."

Response 1: Our evaluation of the possible contribution to children's exposure to food dyes did include a first ever analysis of food dye content in certain children's OTC pharmaceuticals. We draw your attention to findings that have been changed in the final version of the document. We were contacted by Consumer Healthcare Products Association (CHPA), who found an error in the publication we used to estimate exposure to synthetic food dyes through OTC cough/cold/allergy and pain relieve/fever reducer syrups for children. The error is limited to the syrup formulations and did not affect estimates from other formulations. In response to the comments from CHPA, we confirmed the inaccuracies in our initial color additive exposure estimates for children from cough/cold/allergy and pain relieve/fever reducer syrups, and have made changes to the text and tables of the report. With one exception, the corrected values are all 100-times lower than initially reported and **none of the estimated exposure values for OTC syrups exceed the US FDA or JEFCA ADIs.** We believe we have addressed each of CHPA's comments and concerns. Please see the response to comments from the CHPA for details of the changes in the document.

Comment 2: "OEHHA has recommended additional regulatory studies to determine safer, updated ADIs for food dyes. Although more research is always valuable, EWG finds that the existing body of research is sufficient to permit the steps necessary to protect the most sensitive children. Rowe and Rowe observed behavioral effects of irritability, restlessness and sleeplessness in children who were exposed to Yellow No. 5 at levels of just 1 mg.[1] Therefore, it is important to address the immediate need to alert parents to these harms in food products marketed to children and reduce their exposure. Since 2010, food sold in the European Union with any one of six dyes, including Yellow No. 5, Yellow No. 6 and Red. No. 40, must be labeled with the warning "may have an adverse effect on activity and attention in children.""

Response 2: OEHHA makes recommendations for additional research in the conclusions of the document. However, we agree that there are sufficient data now to conclude that the synthetic food dyes can affect children's behavior. We have clarified that in the conclusions (section 8.6) of the report.

Comment 3: Finally, EWG agrees with OEHHA's methodology of using multiple data streams, including high throughput toxicology data, to assess the neurotoxicity of food dyes in children, and we support OEHHA's recommendation of reducing children's exposure to these substances.

Response 3: We have attempted to include and integrate all of the key evidence generated by multiple disciplines. Research has been conducted using a large variety of approaches and outcomes studied over multiple decades of research. We believe that the preponderance of evidence across these studies, combined with our evaluation of exposure, supports our conclusions.

[1] Rowe KS, Rowe KJ. 1994. Synthetic food coloring and behavior: a dose response effect in a double-blind, placebo-controlled, repeated-measures study. The Journal of Pediatrics. 125(5): 691-8. https://doi.org/10.1016/s0022-3476(94)70059-1

Response to Comments from Prof. Jim Stevenson, Emeritus Professor of Developmental Psychopathology, University of Southampton, UK,

Comment 1: "This is a wide ranging and meticulously conducted set of systematic reviews on synthetic food dyes and their impact on children."

"The coverage of the material in Chapter 2 is a very impressive and comprehensive review of the studies undertaken to date. Reference is made to the Stevenson et al. (2010) study identifying polymorphisms in the histamine N-methyl transferase (HNMT) gene which moderated the impact of food dyes on behavior. This is a potentially important indicator of possible pathways via which food dyes can influence behavior. The report correctly suggests that this finding has not been replicated. However, as far as I know, no-one has attempted to replicate it using a RCT challenge of food dyes in children. It should be noted that there is accumulating evidence of the role of HNMT in brain functions (Yoshikawa et al. (2019). Histamine N-transferase in the brain. International Journal of Molecular Sciences, 20, (3): **DOI:** 10.3390/ijms20030737)."

Response 1: There is an increased recognition that histamine is an important neurotransmitter in the brain. As reviewed in Scammel et al. (2019), histaminergic neurons are located in the tuberomaxillary nucleus and innervate other regions of the brain including the cortex, thalamus, and other wake-promoting areas. Histaminergic neurons modulate the function of neurons that make dopamine, serotonin, glutamates, GABA, norepinephrine, and acetylcholine. Thus, histamine can have a wide range of effects on brain function. Histamine and serotonin systems are intertwined biologically (Best et al. 2017), providing a link between the Stevenson study in humans and serotonin mechanism studies in animals that we reviewed.

Histamine N-methyl transferase methylates histamine in the extracellular space following release resulting in inactivation of histamine. A number of polymorphisms in the gene have been identified leading to either greater or lesser activity of the enzyme, and several have been associated with specific neurological diseases (Yoshikawa et al, 2019). Thus, the finding by Stevenson et al. (2010) that HNMT polymorphisms modulate the effect of food dyes on activity adds to the evidence that the food dyes can impact brain function, and in this case behavior.

Comment 2: "I agree with and wish to underscore the statements in Chapter 5 Hazard Identification that ADHD is considered to exist on a spectrum of neurobehavioral symptoms and severity. The evidence from the Bateman et al. (2004) and McCann et al (2007) studies is that the presence of dyes in the diet increases the population mean on measures of hyperactivity by about one fifth of a standard deviations (effect size = 0.20). This is very similar to the effect size reported for high quality studies in the Nigg et al. (2012) meta-analysis (effect size = 0.23). "

Response 2: OEHHA tries to make the point that, considering widespread exposure in the population, these relatively small changes can result in significant impacts on children's behavior, which in turn impact their school performance and other social interactions..

Comment 3: "This 0.20 effect size needs to placed in a public health context. It is very close the effect of environmental lead on children's IQ. Grosse et al. (2002) [Environmental Health Perspectives, 110, 563-569] concluded that ""These calculations imply that, because of falling Blood Lead Levels, U.S. preschool-aged children in the late 1990s had IQs that were, on average, 2.2–4.7 points higher than they would have been if they had the blood lead distribution

observed among U.S. preschool-aged children in the late 1970s ". Taking a mid-point effect of 3.5 IQ points and given the standard deviation of IQ is 15, this produced an effect size of the gain from the reduction of lead of 3.5/15 = 0.23. This suggests that the gains for children's behavior from the removal of dyes would be equivalent to the benefit obtain for IQ of the reduction of lead exposure.

It is important to recognise that this 0.20 is only an average figure – some children's hyperactivity scores will increase more than this, other less so. A feature of this variation in the response to dyes is that it is normally distributed. Below are appended unpublished graphs showing these distributions from the McCann et al (2007) study. [These graphs are available in the original comment also posted on OEHHA's web page.] There is no evidence of a sub-group that are much more severely affected than others. This makes it dubious to attempt to identify the number of children affect by dyes because this is not an all or none effect. It is a question of degree. When we examined factors that might influence the degree of change in Bateman et al. (2004) we found no effects on the degree of behaviour change under dye exposure of initial hyperactivity level or the presence of atopy (allergic sensitivity). In McCann et al. (2007) we found no effect on the degree of change of gender, pretrial diet, mothers education or social class. The only factors influencing vulnerability that we were able to identify were the HNMT polymorphisms reported in Stevenson et al. (2010). There is then no readily available method of identifying which children are most affected.

Exposures to food dyes can be seen as shifting more children towards more severe symptoms. The effect of this shift is enhanced at the high end of the distribution, thereby increasing the numbers of those who meet the criteria for the clinical diagnosis of ADHD, resulting in large costs for society.

Response 3: We agree that the effect size represents a central tendency response, and that some children are more impacted than others. There has not, however, been specific population characteristics for susceptibility identified other than potentially those with specific histamine metabolism enzyme polymorphisms. We have also pointed out the importance of the findings that the effects associated with synthetic food dyes are found across the population tested and not just in children with identified symptoms or diagnoses.

Comment 4: I fully support the conclusions in the final paragraph of section "8.1 Summary of human studies." This summary is balanced and gives due appreciation of the challenges in conducting research on human subjects. It also rightly emphasises the paramount value of findings from well-conducted randomised control trials on this topic.

On p.286 it is recognised that "Research is generally a long-term proposition." This is true. Adequately powered RCTs on children from the general population are time consuming and expensive. Their complexity is markedly increased if they are designed to test the effects of combinations of dyes in mixtures. Whatever further RCTs on the topic are commissioned, it must be acknowledged that there is already evidence that "neurobehavioral effects of synthetic food dyes" are found in children. Harm can be prevented by reduction in exposure to food dyes before further studies are completed. The harm includes not only the immediate symptoms but also later educational difficulties and antisocial behaviour, which can have lifetime consequences - as is shown by the following finding: "There were strong linear relationships between early hyperactivity and later adverse outcomes. Adjustment for other childhood variables suggested that early hyperactivity was associated with continuing school difficulties,

problems with attention and poor reading in adolescence." (McGee et al. (2002) Journal of Child Psychology and Psychiatry, 43, 1004-1017).

The following final conclusion/recommendation is entirely justified:

"At a minimum, in the short-term, the neurobehavioral effects of synthetic food dyes in children should be acknowledged and steps taken to reduce exposure to these dyes in children."

Response 4: Thank you for supplying the reference and quote from McGee et al. (2002). We agree that, while further research would help clarify remaining questions, there is enough evidence now, as noted in our conclusions, that support an association between synthetic food dye consumption and adverse neurobehavioral impacts in children.

Response to Comments from T.J. Sobotka

Comment 1: "In Chapter 2 of this OEHHA draft review document, the authors' presented their conclusions (Section 2.8, p. 45) from the epidemiologic studies of synthetic food dyes and neurobehavioral outcomes in children. Even though the authors stated that "… the findings in the clinical studies they reviewed are not entirely consistent from one study to the next" and that "..Clear associations were not seen in every study, and not all of the studies were high quality", they concluded that "Based on the extent of the positive findings reported, and the fact that we could not convincingly or consistently attribute these positive findings to errors in study design or other bias, we conclude that the current human epidemiologic evidence supports a relationship between food dye exposure and adverse behavioral outcomes in children." In view of the noted inconsistencies across studies and findings of questionable significance from studies of varying quality, the available evidence seems to provide little, if any, consistent or verifiable support for a clear association between food dye exposure and adverse behavioral outcomes in children. Accordingly, the statement of conclusion should be qualified to reflect that the available evidence may suggest or support a possible relationship between certain food dyes and adverse behavioral outcomes in some children."

Response 1: We conducted extensive analyses, described in Chapter 2, to identify biases or other factors that would cast doubt on the positive associations reported in the literature. The conclusion is based on clinical trial studies with a cross-over design that were placebo controlled and blinded. These study design features reduce common biases such as exposure misclassification as well as potential confounding (because the study subjects are compared to themselves). After these extensive analyses, we could not identify any clear biases or confounding factors that would account for the positive associations reported in the clinical trials. Hence, our conclusion that the current human epidemiologic evidence supports a relationship between food dye exposure and adverse behavioral outcomes in some children.

Comment 2: "There is a relevant body of evidence from epidemiologic studies using oligoantigenic or 'few foods' elimination type diets, which was not considered in this OEHHA assessment document. These studies show that certain children with ADHD or other problem behaviors may exhibit a unique intolerance to a variety of foods and food components, including but not limited to artificial colors (e.g., Boris et al, 1994; Carter et al, 1994; Egger et al, 1985; Egger et al, 1992; Kaplan et al, 1989; Schmidt et al, 1997; Uhlig et al, 1997). These studies support the possibility that certain children, including those with ADHD or other problem behaviors, may exhibit a unique intolerance to a variety of foods and food components, including artificial colors. Submitted 9/11/2020 References Boris M and Mandel FS. Foods and additives are common causes of the attention deficit hyperactive disorder in children. Ann Allergy, 1994 (May), 72(5):462-8. Carter CM, Urbanowicz M, Hemsley R, Mantilla L, Strobel S, Graham PJ, Taylor E. Effects of a few food diet in attention deficit disorder. Arch Dis Child, 1993 (November), 69(5):564-8. Egger J, Carter CM, Graham PJ, Gumley D and Soothill JF. Controlled trial of oligoantigenic treatment in the hyperkinetic syndrome. Lancet. 1985 (Mar 9), 1(8428):540-5 Egger J., Stolla A and McEwen LM. Controlled trial of hyposensitization in children with food-induced hyperkinetic syndrome. The Lancet, 1992 (May), 339: 1150-1153. Kaplan BJ, McNicol J, Conte RA and Moghadam HK. Dietary replacement in preschool-aged hyperactive boys. Pediatrics, 1989 (Jan), 83 (No.1): 7-17. Schmidt MH, Mocks P, Lay B, Eisert HG, Fojkar R, Fritz-Sigmund D, Marcus A and Musaeus B. Does oligoantigenic diet influence

hyperactive/conduct-disordered children - a controlled trial. European Child & Adolescent Psychiatry, 1997, 6: 88-95. Uhlig T Merkenschlager A Brandmaier R Egger J. Topographic mapping of brain electrical activity in children with food-induced attention deficit hyperkinetic disorder. Eur J Pediatr, 1997 (July), 156(7):557-61.

Response 2: The focus of OEHHA's report was on synthetic food dyes, not on the effects of broad categories of foods. As noted in our responses to comments from IACM (see responses 1.13), any effects seen in studies involving broad categories of foods could potentially be due to any of a number of different chemical agents in these foods. In contrast, the studies OEHHA included in its review were focused primarily on synthetic food dyes, which is the subject of our report. As such, any effects seen in these studies are much more likely to be due to synthetic food dyes than to any other chemical agent in food.

References Used in OEHHA Responses

Albasher G, Maashi N, Alfarraj S, Almeer R, Albrahim T, Alotibi F, Bin-Jumah M, Mahmoud AM. 2019. Perinatal Exposure to Tartrazine Triggers Oxidative Stress and Neurobehavioral Alterations in Mice Offspring. Antioxidants. 9(1):53.

Amrhein V, Greenland S, McShane B. 2019 Scientists rise up against statistical significance. Nature. 567(7748):305-307.

Asalman N, Aljaferi, A, Wani TA, Zagar S. (2019) High-dose aspirin reverses tartrazine-induced cell growth dysregulation independent of p53 signalling and antioxidant mechanisms in rt brain. Hindawi BioMed Research International.

Axon A., May F, Gaughan LE, Williams FW, Blain PG, Wright MC. 2012. Tartrazine and sunset yellow are xenoestrogens in a new screening assay to identify modulators of human oestrogen receptor transcriptional activity. Toxicology. 298(1-3), 40-51.

Bastaki M, Farrell T, Bhusari S, Bi X, Scrafford C. 2017. Estimated daily intake and safety of FD&C food-colour additives in the us population. Food Additives & Contaminants: Part A 34:891-904

Best J, Nijhout HF, Samaranayake S, Hashemi P, Reed M. A mathematical model for histamine synthesis, release, and control in varicosities. Theor Biol Med Model. 2017 Dec 12;14(1):24.

Bhatt D, Vyas K., Singh S, John PJ, Soni I. 2018. Tartrazine induced neurobiochemical alterations in rat brain sub-regions. Food Chem Toxicol. 113, 322-327.

Borzelleca JF, Hogan GK. 1985. Chronic toxicity/carcinogenicity study of FD & C blue no. 2 in mice. Food Chem Toxicol. 23:719-722.

Borzelleca JF, Hogan GK, Koestner A. 1985. Chronic toxicity/carcinogenicity study of FD & C Blue No. 2 in rats. Food Chem Toxicol. 23:551-558.

Borzelleca, J.F, Capen, C.C, & Hallagan, J.B. 1987a. Lifetime toxicity/carcinogenicity study of FD & C Red No. 3 (erythrosine) in rats. Food and Chemical Toxicology, 25(10), 723–733.

Borzelleca, J.F, & Hallagan, J.B. 1987b. Lifetime toxicity/carcinogenicity study of FD & C Red No. 3 (erythrosine) in mice. Food and Chemical Toxicology. 25(10), 735–737.

Borzelleca JF, Hallagan JB. 1992. Safety and regulatory status of food, drug, and cosmetic color additives. In: Food safety assessment. 484: 377-390.

Bradford Hill A. 1965. The environment and disease: Association or causation? Proceedings of the Royal Society of Medicine. 58:295-300.

Brown DL. 2017. Practical stereology applications for the pathologist. Vet Pathol 54(3):358-368.

Butterworth KR, Hoosen J, Gaunt IF, Kiss IS, Grasso P. 1975. Long-term toxicity of indigo carmine in mice. Food Cosmet Toxicol 13:167-176.

Capen CC. 1998. Correlation of mechanistic data and histopathology in the evaluation of selected toxic endpoints of the endocrine system. Toxicol Lett. 102-103:405-409

Chappell, GA, Britt JK., Borghoff SJ. 2020. Systematic assessment of mechanistic data for FDA-certified food colors and neurodevelopmental processes. Food Chem Toxico. 140.

Chen YH, Tseng CP, How SC, Lo CH, Chou WL, Wang SS. 2016. Amyloid fibrillogenesis of lysozyme is suppressed by a food additive brilliant blue FCF. Colloids Surf B Biointerfaces. 142:351-359.

Choudhary M, Malek G. 2020. The Aryl Hydrocarbon Receptor: A Mediator and Potential Therapeutic Target for Ocular and Non-Ocular Neurodegenerative Diseases. Int J Mol Sci. 21 (18):6777.

Chung KT, Cerniglia CE. 1992. Mutagenicity of azo dyes: Structure-activity relationships. Mutat Res. 277:201-220.

Collins, T.F.X, & Long, E.L. (1976). Effects of chronic oral administration of erythrosine in the Mongolian gerbil. Food and Cosmetics Toxicology, 14(4), 233–248.

Collins TFX, Black TN, Ruggles DI. Teratogenic Potential of FD&C RED NO. 3 When given by Gavage. Toxicology and Industrial Health. 1993;9(4):605-616.

Collins, T.F.X, Black, T.N, O'Donnell, M.W, Shackelford, M.E, & Bulhack, P. (1993). Teratogenic potential of FD & C Red No. 3 when given in drinking water. Food and Chemical Toxicology, 31(3), 161–167.

Cummings, Steven R., Hulley, Stephen B., Newman, Thomas B., Browner, Warren S., Grady, Deborah G. Designing Clinical Research. United States: Wolters Kluwer/Lippincott Williams & Wilkins, 2015.

Dalal A, Poddar MK. 2009. Short-term erythrosine B-induced inhibition of the brain regional serotonergic activity suppresses motor activity (exploratory behavior) of young adult mammals. Pharmacol Biochem Behav. 92:574-582.

Dalal A, Poddar MK. 2010. Involvement of high plasma corticosterone status and activation of brain regional serotonin metabolism in long-term erythrosine-induced rearing motor hyper activity in young adult male rats. Toxicol Mech Methods. 20:287-297.

Davis AP, Grondin CJ, Johnson RJ, Sciaky D, McMorran R, Wiegers J. 2019. The comparative toxicogenomics database: Update 2019. Nucleic Acids Res 47:D948-D954.

Dees C, Askari M, Garrett S, Gehrs K., Henley D, Ardies CM. 1997. Estrogenic and DNA-damaging activity of Red No. 3 in human breast cancer cells. Environ Health Perspect. 105 (Suppl 3), 625–632.

Doell DL, Folmer DE, Lee HS, Butts KM, Carberry SE. 2016. Exposure estimate for FD&C colour additives for the us population. Food Addit Contam Part A Chem Anal Control Expo Risk Assess. 33:782-797

Doguc DK, Ceyhan BM, Ozturk M, Gultekin F. 2013. Effects of maternally exposed colouring food additives on cognitive performance in rats. Toxicol Ind Health. 29:616-623.

Doguc DK, Aylak F, Ilhan I, Kulac E, Gultekin F. 2015. Are there any remarkable effects of prenatal exposure to food colourings on neurobehaviour and learning process in rat offspring? Nutr Neurosci. 18:12-21.

Erickson ZT, Falkenberg EA, Metz GA. 2014. Lifespan psychomotor behaviour profiles of multigenerational prenatal stress and artificial food dye effects in rats. PLoS One 9:e92132.

Erskine HE, Norman RE, Ferrari A, Chan GCK, Copeland WE, Whiteford HA, Scott JG. 2016. Long-Term outcomes of Attention-Deficit/ Hyperactivity Disorder and Conduct Disorder: A systematic review and meta-analysis. J Am Acad Child Adolesc Psychiatry. 55(10):841–850

Feng J, Cerniglia CE, Chen H. 2012. Toxicological significance of azo dye metabolism by human intestinal microbiota. Front Biosci (Elite Ed). 4:568-586.

Ganesan L, Margolles-Clark E, Song Y, Buchwald P. 2011. The food colorant erythrosine is a promiscuous protein-protein interaction inhibitor. Biochem Pharmacol. 81:810-818.

Ganesan L, Buchwald P. 2013. The promiscuous protein binding ability of erythrosine B studied by metachromasy (metachromasia). J Mol Recognit. 26:181-189.

Gao Y, Li C, Shen J, Yin H, An X, Jin H. 2011. Effect of food azo dye tartrazine on learning and memory functions in mice and rats, and the possible mechanisms involved. J Food Sci. 76:T125-T129.

Gaunt IF, Kiss IS, Grasso P, Gangolli SD. 1969. Short-term toxicity study on indigo carmine in the pig. Food Cosmet Toxicol. 7:17-24.

Goldenring JR, Wool RS, Shaywitz BA, Batter DK, Cohen DJ, Young JG. 1980. Effects of continuous gastric infusion of food dyes on developing rat pups. Life Sci. 27:1897-1904

Goyette GH, Connors CK, Petti TA, Curtis LE. 1978. Effects of artificial colors on hyperkinetic children: A double-blind challenge study [proceedings]. Psychopharmacol Bull. 14:39-40.

Guyton CL, Reno FE. Metabolic disposition of 35S Allura Red AC in the dog and rat. Unpublished report no. 165-146 by Hazelton Laboratories America, Inc.; 1975. Submitted to WHO by Allied ChemicalCorporation

Guyton CL, Stanovick RP. 1975. Determination of the metabolites of Allura Red AC in the rat and dog. Unpublished report by Hazelton Laboratories America, Inc. Submitted to WHO by AlliedChemical Corporation.

Hansen WH, Fitzhugh OG, Nelson AA, Davis KJ. 1966a. Chronic toxicity of two food colors, brilliant blue FCF and indigotine. Toxicol Appl Pharmacol. 8:29-36.

Hansen WH, Long EL, Davis KJ, Nelson AA, Fitzhugh OG. 1966b. Chronic toxicity of three food colourings: Guinea green b, light green sf yellowish and fast green FCF in rats, dogs and mice. Food Cosmet Toxicol. 4:389-410.

Harp BP, Miranda-Bermudez E, Barrows JN. 2013. Determination of seven certified color additives in food products using liquid chromatography. J Agric Food Chem. 61(15):3726-3736.

Hollingsworth RL. 1982. Color additive petition no. 8c-0064.

Iyer S, Pham N, Marty M, Sandy M, Solomon G, Zeise L. 2019. An integrated approach using publicly available resources for identifying and characterizing chemicals of potential toxicity concern: Proof-of-concept with chemicals that affect cancer pathways. Toxicol Sci. 169:14-24.

Judson R, Houck K, Martin M, Richard AM, Knudsen TB, Shah I. 2016. Editor's highlight: analysis of the effects of cell stress and cytotoxicity on in vitro assay activity across a diverse chemical and assay space. Toxicol Sci. 152, 323–339.

Karkada G, Shenoy KB, Halahalli H, Karanth KS. 2012. Nardostadys jatamansi extracts prevent chronic restraint stress-induced learning and memory deficits in radial arm maze task. J Natural Sciences, Bilogy and Medicine. 3(2):125-132.

Khayyat L, Essawy A, Sorour J, Soffar A. 2017. Tartrazine induces structural and functional aberrations and genotoxic effects in vivo. PeerJ 5:e3041.

Kisby GE, Spencer PS. Is neurodegenerative disease a long-latency response to early-life genotoxin exposure? Int J Environ Res Public Health. 2011 Oct;8(10):3889-921

Klein RG, Mannuzza S, Ramos Olazagasti MA, Roizen Belsky E, Hutchison JA, Lashua-Shriftman E, Castellanos FX. .2012. Clinical and functional outcome of childhood ADHD 33 years later. Arch Gen Psychiatry. 69(12): 1295–1303.

Klimisch H-J, Andreawe M, Tillman U. 1997. A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data. Regulatory Toxicol Pharmacol. 25:1-5.

Kobylewski S, Jacobson MF. 2012. Toxicology of food dyes. Int J Occup Environ Health. 18:220-246.

Kruhlak NL, Contrera JF, Benz RD, Matthews EJ. 2007. Progress in QSAR toxicity screening of pharmaceutical impurities and other FDA regulated products. Adv Drug Deliv Rev. 59:43–55.

Lau K, McLean WG, Williams DP, Howard CV. 2006. Synergistic interactions between commonly used food additives in a developmental neurotoxicity test. Toxicol Sci. 90:178-187.

Lee J, Kwon I, Jang SS, Cho AE. 2016. Investigation of the effect of erythrosine B on amyloid beta peptide using molecular modeling. J Mol Model. 22.

Levitan H, Ziylan Z, Smith QR, Takasato Y, Rapoport SI. 1984. Brain uptake of a food dye, erythrosin b, prevented by plasma protein binding. Brain Res. 322:131-134.

Levitan H, Ziya Ziylan Y, Rapoport SI. 1985. Brain uptake of the food dye, erythrosin-b. IRCS MED SCI. 13:64-65.

Levy F, Hobbes G. 1978a. Hyperkinesis and diet: A replication study. Am J Psychiatry. 135:1559-1560.

Levy F, Dumbrell S, Hobbes G, Ryan M, Wilton N, Woodhill JM. 1978b. Hyperkinesis and diet: A double-blind crossover trial with a tartrazine challenge. Med J Aust. 1:61-64.

Lok KY, Chan RS, Lee VW, Leung PW, Leung C, Leung J. 2013. Food additives and behavior in 8- to 9-year-old children in hong kong: A randomized, double-blind, placebo-controlled trial. J Dev Behav Pediatr. 34:642-650.

Mailman RB, Ferris RM, Tang FL, Vogel RA, Kilts CD, Lipton MA. 1980. Erythrosine (Red No. 3) and its nonspecific biochemical actions: What relation to behavioral changes? Science. 207:535-537.

Majdak P, Ossyra JR, Ossyra JM, Cobert AJ, Hofmann GC, Tse S, Panozzo B, Grogan EL, Sorokina A, Rhodes JS. 2016. A new mouse model of ADHD for medication development. Sci Rep. 6:39472.

Mailman RB. 1987. Mechanisms of CNS injury in behavioral dysfunction. Neurotoxicol Teratol 9:417-426.

Mailman RB, Lewis MH. 1987. Neurotoxicants and central catecholamine systems. NeuroToxicology. 8(1):123-140.

Mathieu-Denoncourt J, Martyniuk CJ, de Solla SR, Balakrishnan VK, Langlois VS. 2014. Sediment contaminated with the azo dye disperse yellow 7 alters cellular stress- and androgen-related transcription in silurana tropicalis larvae. Environ Sci Technol. 48: 2952-2961.

McCann D, Barrett A, Cooper A, Crumpler D, Dalen L, Grimshaw K. 2007. Food additives and hyperactive behaviour in 3-year-old and 8/9-year-old children in the community: A randomised, double-blinded, placebo-controlled trial. Lancet. 370:1560-1567

McCarthy MM. 2008. Estradiol and the developing brain. Physiol Rev. 88:91-124.

McEwen BS, Akama KT, Spencer-Segal JL, Milner TA, Waters EM. 2012. Estrogen effects on the brain: actions beyond the hypothalamus via novel mechanisms. Behavior Neurosci. 126:4-16.

McGee R, Prior M, Williams S, Smart D, Sandon A. 2002. The long-term significance of teacher-rated hyperactivity and reading ability in childhood: findings from two longitudinal studies. J Child Psych Psyc, 43, 1004-1017

Minegishi K, Morimoto K, Yamaha T. 1978. Metabolism of triphenylmethane colors (iii). Comparison of tissue distribution and biliary excretion for 3h-benzyl violet 4b and 3h-fast green FCF (food green no. 3) in rats. J Food Hyg Soc Jpn. 19:482-485.

Mohamed AAR, Galal AAA, Elewa YHA. 2015. Comparative protective effects of royal jelly and cod liver oil against neurotoxic impact of tartrazine on male rat pups brain. Acta Histochem. 117, 649-658.

Nigg JT, Lewis K, Edinger T, Falk M. 2012. Meta-analysis of attention-deficit/hyperactivity disorder or attention-deficit/hyperactivity disorder symptoms, restriction diet, and synthetic food color additives. J Am Acad Child Adolesc Psychiatry. 51:86-97.e88.

Noorafshan A, Hashemi M, Karbalay-Doust S, Karimi F. 2018. High dose allura red, rather than the adi dose, induces structural and behavioral changes in the medial prefrontal cortex of rats and taurine can protect it. Acta Histochem. 120:586-594

NTP Office of Health Assessment and Translation. 2019. Risk of Bias Tool. National Toxicology Program. Accessed Dec 9, 2019. https://ntp.niehs.nih.gov/whatwestudy/assessments/noncancer/riskbias/index.html.

Obrist J, LeVan L, Puhl R, Duan R. 1986. Final report: Metabolism of FD&C Red No. 3 in rats, study no. 6145–100. Unpublished report by hazleton laboratories america, inc

Osman MY, Sharaf IA, el-Rehim WMA, el-Sharkawi AM. 2002. Synthetic organic hard capsule colouring agents: In vitro effect on human true and pseudo-cholinesterases. Br J Biomed Sci. 59:212-217.

Osman MY, Sharaf IA, Osman HMY, El-Khouly ZA, Ahmed EI. 2004. Synthetic organic food colouring agents and their degraded products: Effects on human and rat cholinesterases. Br J Biomed Sci 61:128-132.

Paul Friedman K, Watt ED, Hornung MW, Hedge JM, Judson RS, Crofton KM, Houck KA, Simmons SO. 2016. Tiered High-Throughput Screening Approach to Identify Thyroperoxidase Inhibitors Within the ToxCast Phase I and II Chemical Libraries. Toxicol Sci.151(1):160-80.

Pelsser LM, Frankena K, Toorman J, Rodrigues Pereira R. 2017. Diet and ADHD, Reviewing the Evidence: A Systematic Review of Meta-Analyses of Double-Blind Placebo-Controlled Trials Evaluating the Efficacy of Diet Interventions on the Behavior of Children with ADHD. PLoS One. 12(1):e0169277.

Pham N, Iyer S, Hackett E, Lock BH, Sandy M, Zeise L, Solomon G, Marty M. 2016. Using ToxCast to explore chemical activities and hazard traits: A case study with ortho-phthalates. Toxicol Sci. 151(2):296-301.

Pollock I, Warner JO. 1990. Effect of artificial food colours on childhood behaviour. Arch Dis Child. 65:74-77.

Pritchett K, Mulder GB. (2003) T, radial arm, and Barnes mazes. Contemp Top Lab Animal Sci. Volume 42(3): 53-55.

Reisen CA, Rothblat LA. 1986. Effect of certified artificial food coloring on learning and activity level in rats. Neurobehav Toxicol Teratol. 8:317-320.

Rafati A, Nourzei N, Karbalay-Doust S, Noorafshan A. 2017. Using vitamin e to prevent the impairment in behavioral test, cell loss and dendrite changes in medial prefrontal cortex induced by tartrazine in rats. Acta Histochem 119:172-180.

Rose TL. 1978. The functional relationship between artificial food colors and hyperactivity. J Appl Behav Anal. 11:439-446.

Rothman, K. J. and S. Greenland. 1998. Modern Epidemiology, 2nd edition. Lippincott-Raven

Rowe KS, Rowe KJ. 1994. Synthetic food coloring and behavior: A dose response effect in a double-blind, placebo-controlled, repeated-measures study. J Pediatr. 125:691-698.

Scammel TE, Jackson AC, Franks NP, Wisden W, Dauvilliers Y. 2019. SLEEPJ, 2019, 1–8.

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

Segal D, Makris SL, Kraft AD, Bale AS, Fox J, Gilbert M, Bergfeldt DR, Raffaele KC, Blain RB, Fedak KM, Selgrade MK, Crofton KM. 2015. Evaluation of the ToxRTool's ability to rate the reliability of toxicological data for human health hazard assessments. Reg Toxicol Pharmacol 72:94-101.

Shaywitz BA, Goldenring JR, Wool RS. 1979. Effects of chronic administration of food colorings on activity levels and cognitive performance in developing rat pups treated with 6-hydroxydopamine. Neurobehav Toxicol 1:41-47.

Silva M, Pham N, Lewis C, Iyer S, Kwok E, Solomon G, Zeise L. 2015. A Comparison of ToxCast Test Results with In Vivo and Other In Vitro Endpoints for Neuro, Endocrine, and Developmental Toxicities: A Case Study Using Endosulfan and Methidathion. Birth Defects Research Part B Developmental and Reproductive Toxicology.

Sobin C, Golub M. 2018. Behavioral Outcome as a Primary Organizing Principle for Mechanistic Data in Developmental Neurotoxicity; Chapter 29 in: Handbook of Developmental Neurotoxicology.

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

Sonuga-Barke EJ, Brandeis D, Cortese S, Daley D, Ferrin M, Holtmann M, Stevenson J, Danckaerts M, van der Oord S, Döpfner M, Dittmann RW, Simonoff E, Zuddas A, Banaschewski T, Buitelaar J, Coghill D, Hollis C, Konofal E, Lecendreux M, Wong IC, Sergeant J. 2013. European ADHD Guidelines Group. Nonpharmacological interventions for ADHD: systematic review and meta-analyses of randomized controlled trials of dietary and psychological treatments. Am J Psychiatry. 170(3):275-89.

Stevenson J, Sonuga-Barke E, McCann D, Grimshaw K, Parker KM, Rose-Zerilli MJ. 2010. The role of histamine degradation gene polymorphisms in moderating the effects of food additives on children's adhd symptoms. Am J Psychiatry. 167:1108-1115.

Stevenson J, Buitelaar J, Cortese S, Ferrin M, Konofal E, Lecendreux M, Simonoff E, Wong IC, Sonuga-Barke E. 2014. Research review: the role of diet in the treatment of attention-

deficit/hyperactivity disorder--an appraisal of the evidence on efficacy and recommendations on the design of future studies. J Child Psychol Psychiatry. 55(5):416-27.

Tanaka T. 1994. Reproductive and neurobehavioral effects of allura red ac administered to mice in the diet. Toxicology. 92:169-177.

Tanaka T. 1996. Reproductive and neurobehavioral effects of sunset yellow FCF administered to mice in the diet. Toxicol Ind Health 12:69-79.

Tanaka T. 2001. Reproductive and neurobehavioural toxicity study of erythrosine administered to mice in the diet. Food Chem Toxicol 39:447-454.

Tanaka T. 2006. Reproductive and neurobehavioural toxicity study of tartrazine administered to mice in the diet. Food Chem Toxicol 44:179-187.

Tanaka T, Takahashi O, Oishi S, Ogata A. 2008. Effects of tartrazine on exploratory behavior in a three-generation toxicity study in mice. Reprod Toxicol 26:156-163.

Tanaka T, Takahashi O, Inomata A, Ogata A, Nakae D. 2012. Reproductive and neurobehavioral effects of brilliant blue FCF in mice. Birth Defects Res B Dev Reprod Toxicol 95:395-409.

Tanaka T. 2015b. Comparison of measurements of the same variables of exploratory behavior in mice with different apparatuses. J Exp Appl Anim Sci. 1(3):301-316.

Tanaka T. 2015a. Sex differences in exploratory behavior of laboratory CD-1 mice (mus musclus). Scandin J Lab Anim Sci. 41(5) ISSN 2002-0112

Tanaka T, Ogata A, Inomata A, Nakae D. 2014. Effects of different types of bedding materials on behaioral development in laboratory CD1 mic (Mus musculus). Birth Defects Res B Reprod Toxicol. 101:393-401.

Truong L, Reif DM, St Mary L, Geier MC, Truong HD, Tanguay RL. 2014 Multidimensional in vivo hazard assessment using zebrafish. Toxicol Sci. 137 (1):212-233.

US EPA. 1986. Guidelines for the Health Risk Assessment of Chemical Mixtures. U.S. Environmental Protection Agency, Washington D.C. EPA/630/R-98/022.

US EPA (2000) Supplementary Guidance for Conducting Health Risk Assessments of chemical Mixtures. U.S. Environmental Protection Agency, Washington, D.C. EPA 630/R-00/002 August, 2000

US FDA. (2011) Background document for the food advisory committee: Certified color additives in food and possible association with attention deficit hyperactivity disorder in children march 30-31, 2011.

US FDA Redbook (2000) Guidance for Industry and Other Stakeholders. Toxicological Principles for the Safety Assessment of Food Ingredients. U.S. Department of Health and

Human Services, Food and Drug Administration, Center for Food Safety and Applied Nutrition, July 2000; Updated July 2007.

Vorhees C, Butcher R, Brunner R, Wootten V, Sobotka T. 1983a. A developmental toxicity and psychotoxicity evaluation of fd and c red dye #3 (erythrosine) in rats. Arch Toxicol. 53:253-264.

Vorhees C, Butcher R, Brunner R, Wootten V, Sobotka T. 1983b. Developmental toxicity and psychotoxicity of fd and c red dye no. 40 (allura red ac) in rats. Toxicology. 28:207-217.

Vyas S, Rodrigues, AJ, Silva JM. 2016. Chronic Stress and Glucocorticoids: From Neuronal Plasticity to Neurodegeneration. Neural Plast. 6391686.

Wenk GL. 2004. Assessment of Spatial Memory Using the UNIT 8.5A Radial Arm Maze and Morris Water Maze. Current Protocols in Neuroscience. 8.5A.1-8.5A.12

White RG. Metabolic fate of orally ingested non-toxic Red Z-4576. Unpublished report no. 21855 by Buffalo Research Laboratory; 1970. Submitted to WHO by Allied Chemical Corporation.

Williams AJ, Grulke CM, Edwards J, McEachran AD, Mansouri K, Baker NC. 2017. The comptox chemistry dashboard: A community data resource for environmental chemistry. Journal of Cheminformatics 9:61.

Wójtowicz AK, Szychowski KA, Wnuk A, Kajta M. 2017. Dibutyl Phthalate (DBP)-Induced Apoptosis and Neurotoxicity are Mediated via the Aryl Hydrocarbon Receptor (AhR) but not by Estrogen Receptor Alpha (ERα), Estrogen Receptor Beta (ERβ), or Peroxisome Proliferator-Activated Receptor Gamma (PPARγ) in Mouse Cortical Neurons. Neurotox Res. 31(1):77-89.

Wong HE, Kwon I. 2011. Xanthene food dye, as a modulator of alzheimer's disease amyloid-beta peptide aggregation and the associated impaired neuronal cell function. PLoS One 6.

Wu PY, Chuang PY, Chang GD, Chan YY, Tsai TC, Wang BJ, Lin KH, Hsu WM, Liao YF, Lee H. 2019. Novel Endogenous Ligands of Aryl Hydrocarbon Receptor Mediate Neural Development and Differentiation of Neuroblastoma. ACS Chem Neurosci. 10(9):4031-4042.

Yoshikawa T, Nakamura T, Yanai K. 2019. Histamine N-methyltransferase in the brain. Int. J. Mol. Sci. 2019, 20, 737.