

PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Consumption of ultra-processed foods and associated sociodemographic factors in the US between 2007 and 2012: evidence from a nationally representative cross-sectional study
AUTHORS	BARALDI, LARISSA; Martinez Steele, Euridice; Canella, Daniela; Monteiro, Carlos

VERSION 1 – REVIEW

REVIEWER	Aydin Nazmi California Polytechnic State University, San Luis Obispo, USA
REVIEW RETURNED	01-Dec-2017

GENERAL COMMENTS	<p>Methods</p> <p>Line 108: omit 'and' and 'the'</p> <p>Authors use 'American White' 'American Black', but this is not consistent throughout the manuscript, and unconventional in general use. Suggest standardizing throughout using race/ethnicity terminology per NHANES. https://www.cdc.gov/nchs/nhanes/2011-2012/demo_g.htm#RIDRETH1</p> <p>Assumptions: Categorizing foods into NOVA categories based on NHANES, which was not intended to facilitate such categorization, is complex. I might suggest, perhaps in the Discussion or Methods, listing some of the assumptions inherent in classifying, for example, 'cakes/cookies/pies' as 'ultra-processed', whereas it is reasonable possible that some homemade/bakery cakes may actually be in the 'processed' category. This might also be applicable to other UP foods. How did authors deal with this type of uncertainty? In how many categories/foods could this lead to classification bias?</p> <p>In lines 182-186, authors explain how they dealt with 'handmade' recipes, classifying underlying ingredients. In an average homemade meal that might fall into NOVA category 'processed' due to its incorporation of 'unprocessed' plus 'culinary ingredients', for example, would the approach taken by authors underestimate energy in the 'processed' category? For instance, would a homemade lasagna meal, instead of being categorized as 'processed', actually be classified as 'minimally processed' and 'culinary ingredients'? Please clarify.</p>
-------------------------	---

	<p>Results</p> <p>Tables 1 and 2: would be helpful to add SD to means</p> <p>Table 2: footnotes are complex to interpret. Notably, * and ** seem to be inconsistently applied. Are they needed?</p> <p>Line 258: should be 'inversely'.</p> <p>Line 259: do authors have an exact p-value (ie instead of listing <0.05)?</p> <p>Sentence on lines 264-266: Suggest omitting 'both together', replacing with something akin to: 'However, differences in ultra-processed food consumption as a proportion of total energy between non-Hispanic Black and White groups disappeared, whereas consumption in both groups remained significantly higher compared to other race/ethnic groups'.</p> <p>Discussion</p> <p>Might suggest comparing these NOVA findings with other NHANES studies in terms of: % energy from macronutrients, soft drinks, or other foods based on available data. This would help contextualize these findings with conventional methods of dietary assessment. For example, in lines 230-240 of Results, authors list distributions of energy intake according to NOVA- and in the Discussion, it might be helpful to see what other analyses not using NOVA have found. Authors can then compare, contrast, and speculate on reasons for the differences.</p> <p>Line 319-320: "culinary recipes" is unclear, should this be "homemade foods/meals" or something similar? "turning them into an important eating pattern" is similarly unclear, I think because 'them' is improperly used.</p> <p>Line 321: 'less' should be 'least'. Might suggest 'individuals with a college education consumed the least...' and not using the 'exposed' language, as NHANES measured consumption, not 'exposure'</p> <p>In other instances, authors should update to reflect 'individuals with high school education' and not 'high school level individuals', eg lines 326-327.</p> <p>Lines 365-370: Sentence beginning "As a matter of fact..." is unclear. Perhaps something like, "Indeed, in recent years, patterns of food intake among children and adolescents have become less healthy"?</p> <p>Line 367-368: There is a subject-verb/grammar issue that makes this sentence unclear. Revise. Who/what had the greatest caloric increase? The children or the foods? Also: what were the 'ready-made Mexican dishes'?</p> <p>Line 370: 'fast foods' instead of 'fast food restaurants'</p> <p>Line 371: suggest beginning sentence with 'Our findings are consistent with findings that...'</p>
--	--

	<p>Line 383: in an attempt to preserve cultural identity, yes, but also perhaps out of simple routine and not having been fully acculturated to the fast food culture yet.</p> <p>Line 390-391: suggest for clarity 'may lead individuals to value nutrition over...'</p> <p>Line 397: replace 'breaks up' with 'counters'</p> <p>Line 405: omit 'Even though', begin sentence with "Socioeconomic disparities..." and begin new sentence with 'Our study suggests' (not shows)</p> <p>Line 408: suggest replacing 'gaps' with 'inequities' Also: revise to read 'high overall'</p> <p>Line 409: replace 'superior' with 'higher'</p> <p>Line 410: replace 'fomented' with 'encouraged'</p>
--	---

REVIEWER	Jean Adams Centre for Diet & Activity Research, University of Cambridge, UK
REVIEW RETURNED	06-Dec-2017

GENERAL COMMENTS	<p>Peer review of "Consumption of ultra-processed foods and associated sociodemographic factors in the US between 2007 and 2012"; Larissa Baraldi et al; BMJ Open</p> <p>This is a descriptive study of socio-demographic correlates of ultra-processed food consumption in the USA. It adds to a growing body of such studies from around the world. Whilst useful in a descriptive sense, I would have preferred to see consideration of potential interactions between co-variables considered, as well as more discussion of the public health implications of the results reported.</p> <p>Line 61. Is there evidence specifically of an exponential increase in other countries (ie has an exponential function been tested?), or do you just mean it's increasing at a rapid rate?</p> <p>Line 63-7 (and similar in discussion). It seems possible that there is differential social desirability bias across socio-economic groups.</p> <p>Line 110. I think you mean complicated, not complex? Complex implies unpredictable and I suspect NHANES sampling is entirely predictable if complicated!</p> <p>Lines 120-5. Percentage response rates would be helpful in this paragraph.</p> <p>Line 126. How were data collected from young children?</p> <p>Line 127. Why were pregnant women and breastfeeding mothers excluded? Is there a reason they would be likely to consume a greater percentage of energy from UPF than others?</p> <p>Line 142. What happens in households where there is no household reference person >24 years?</p> <p>Line 182. I was curious to know how you determine if foods are handmade. I followed up the reference cited (Steele et al.; 2016) and this provides no further explanation. Please can you add this.</p> <p>Line 183. I don't see any information on how portion size was estimated. Please add this.</p> <p>Line 210-5. I thought it might be helpful to formally explore differential effects of time across socio-demographic variables using interaction tests.</p>
-------------------------	--

	<p>Line 222-4. To minimise chance findings from multiple comparisons you should reduce to the statistical significance level to 0.01? Two-tailed 0.05 is standard.</p> <p>Line 256-7. Please clarify what the comparator is when stating eg 'higher in adolescents and American Black'.</p> <p>Table 2. I would have liked to have seen some estimate of precision, such as 99% confidence intervals for the values presented in this table – particularly the adjusted ones.</p> <p>Line 302. I'd prefer to see the adjusted figures in the main text (rather than unadjusted), or at least in supplement data.</p> <p>Line 330-44. It would be nice to see some proposed explanations for the differences reported in this paragraph compared to the current results. Why might socio-economic gradients in UPF consumption be different in different countries? Why is UPF consumption higher in the USA even compared to Canada and the UK?</p> <p>Line 345-56. Given that UPF are heterogeneous, I wondered whether the apparent difference between time trends in AHEI and UPF consumption may reflect differences in what UPF are consumed. This is a wider point that relates to the whole discussion – is it possible to choose more vs less 'healthy' UPF? If so, might this partly explain some of the e.g. relatively shallow socio-economic gradients – more affluent people might be choosing a 'healthier' mix of UPF?</p> <p>Discussion. Throughout the discussion there is a tendency to compare the results to previous analyses of diet 'quality' in NHANES using a variety of other markers. This seems to assume that UPF consumption is just another marker of diet quality – is that the case? If so, I think a more robust discussion of what makes it better/worse than existing measures of diet quality would be helpful. If you are proposing UPF consumption as importantly different from other measures of diet quality, perhaps that could also be made more clear.</p> <p>Line 393-404. You quote a newspaper article to claim that 'junk food' is not necessarily cheaper than 'real food'. There is abundant peer-reviewed evidence on this topic that would be preferable to cite.</p> <p>Line 410-12. Please provide evidence to support your claim that regulating marketing and promoting healthier food environments are the most effective way to change UPF consumption.</p> <p>Line 426. Are region and urban/rural measured in NHANES? If so, could they not be included in your analyses?</p> <p>Discussion. I'd like to see some consideration of the how e.g. 60% of energy from UPF should be interpreted from a clinical/public health point of view. Is this a 'healthy' amount? If not, what level should individuals and populations be aiming for? Similarly, it would be helpful to understand how important the statistically significant differences reported are. What difference in disease risk between e.g. 57.9% energy from UPF and 60.4% might be expected?</p> <p>Ethics statement. Whilst the current analysis did not required ethics approval, I'm sure NHANES did. For completeness, it would be helpful to provide details of this.</p> <p>Data sharing. Whilst the authors cannot share data, the data used are presumably available to other researchers on reasonable request to NHAMES. This should be clarified.</p> <p>Checklist. I think the appropriate reporting checklist would be STROBE-Nut, not STROBE.</p>
--	---

VERSION 1 – AUTHOR RESPONSE

Reviewer 1

Line 61. Is there evidence specifically of an exponential increase in other countries (ie has an exponential function been tested?), or do you just mean it is increasing at a rapid rate?

Authors: In fact, no exponential function has been tested; we just meant that it is increasing at a rapid rate. In order to clarify this, we have rephrased this sentence as follows:

“Unlike most articles which have focused on specific food items such as soft drinks or fast food, our study evaluates the impact of a comprehensive group of products whose consumption is increasing rapidly in most countries.”

Line 63-7 (and similar in discussion). It seems possible that there is differential social desirability bias across socio-economic groups.

Authors: This is a good point. We have now included this limitation in both sections as follows:

Strengths and limitations of this study/ Discussion

Social desirability bias may lead to underestimation of ultra-processed food consumption. Should this underreporting have increased with time in response to a growing awareness of the health effects of ultra-processed foods, this could result in a greater underestimation of ultra-processed food consumption in later years. Differential social desirability bias across socio-economic groups (should this exist) could lead to both under or overestimation of the studied associations.

Line 110. I think you mean complicated, not complex? Complex implies unpredictable and I suspect NHANES sampling is entirely predictable if complicated!

Authors: We used the term “complex” as a statistical term for a type of sampling design. The most defining feature of a complex sample is that sample members do not have equal probability of being selected.

Lines 120-5. Percentage response rates would be helpful in this paragraph.

Authors: We accepted this suggestion and added this information as follow:

Of the 10,149 people screened in NHANES 2007–2008, 9,255 (91.2%) participated in the dietary interview in the MEC and 7,838 (77.2%) answered the follow up dietary recall. Similarly, in the second NHANES cycle (2009-2010) of 10,537 people screened, 9,754 (92.6%) and 8,406 (79.8%) completed the first and the follow up dietary recall, respectively. In 2011-12, of the 9,756 individuals screened, 8,519 (87.3%) responded to the first dietary recall and 7,605 (78.%) completed the second dietary recall.

Line 126. How were data collected from young children?

Authors: Thank you for raising this issue. We have added this information as follows: “For children under 9 years of age, the interview was conducted with a proxy; for children between 6 and 8 years of age, in the presence of the child. Children aged 9–11 years provided their own data assisted by an adult household member (assistant). The preferred proxy/assistant was the most knowledgeable person about the child's consumption on the day before the interview. If the child had more than one caregiver, several individuals could contribute to the intake data.” (line 126)

Line 127. Why were pregnant women and breastfeeding mothers excluded? Is there a reason they would be likely to consume a greater percentage of energy from UPF than others?

Authors: Pregnant and breastfeeding women may follow special diets which could lead to spurious associations between SES and ultra-processed food consumption.

Line 142. What happens in households where there is no household reference person >24 years?

Authors: We have now described this in the manuscript as follows:

There were 113 households (1.51% of the sample) where the reference person was younger than 25 years old. In these households, the education achievement of the interviewee (in all cases aged between 18 and 24 years) was adopted.

Line 182. I was curious to know how you determine if foods are handmade. I followed up the reference cited (Steele et al.; 2016) and this provides no further explanation. Please can you add this.
Authors: Thank you for your observation. We have further explained our criteria in the manuscript as follows:

For all food items (Food codes) judged to be a handmade recipe, the classification was applied to the underlying ingredients (Standard Reference codes or SR codes) obtained from the United States Department of Agriculture (USDA) Food and Nutrient Database for Dietary Studies (FNDDS 4.0, 5.0 and 6.6). For example, for cakes, cookies or pies underlying SR Codes were used unless underlying SR Codes or ingredients were unlikely to be used in home recipe (i.e. "Cellulose (Alpha-cellulose, powdered cellulose, and poly-cellulose)" or "Oil, industrial, soy (partially hydrogenated), multiuse for non-dairy butter flavor", or "Shortening, industrial, soybean (hydrogenated) and cottonseed", or "Whey, sweet, dried").

Other examples are "Salsa, red, cooked, not homemade" and "Salsa, red, cooked, homemade". Food Code "Salsa, red, cooked, not homemade" was classified as ultra-processed based on information from a similar product included in the Food code website ("Red Gold Salsa, Mild": tomato concentrate (water, tomato paste), diced tomatoes, jalapeno peppers, green chiles, yellow chiles, vinegar, salt, dried onion, dried garlic, cilantro, natural flavor). On the other hand, for Food Code "Salsa, red, cooked, homemade" it was each underlying SR Code (Salt, table; Garlic, raw; Onions, raw; Tomatoes, red, ripe, canned, packed in tomato juice; Peppers, hot chile, sun-dried; Water, tap, drinking; Vegetable oil, NFS) instead, that was classified according to NOVA, as further explained in a previously published paper. [1]

Line 183. I don't see any information on how portion size was estimated. Please add this.

Authors: Thank you for your observation. We have added the following manuscript (line 121):
"Portion sizes were estimated by NHANES as further explained in NHANES Manual (CDC,2017). Shortly, each MEC dietary interview room contains a standard set of measuring guides that are used to help the respondent report the volume and dimensions of the food items consumed. Upon completion of the in-person interview, participants are given measuring cups, spoons, a ruler, and a food model booklet to use for reporting food amounts during the telephone interview."

Line 210-5. I thought it might be helpful to formally explore differential effects of time across sociodemographic variables using interaction tests.

Authors: Thank you for your suggestion. We have tested the differential effects of time across each sociodemographic variables using interaction terms and added this information in the Results, as follows:

"Significant interaction terms were found for gender-year and age-year ($p < 0.05$)."

Line 222-4. To minimize chance findings from multiple comparisons you should reduce to the statistical significance level to 0.01? Two-tailed 0.05 is standard.

Authors: Thank you for your observation. We agree that two-tailed 0.05 is standard and have now removed "To minimize chance findings from multiple comparisons" from the phrase.

Line 256-7. Please clarify what the comparator is when stating e.g. 'higher in adolescents and American Black'.

Authors: Thank you for highlighting this. We have modified the complete paragraph above as follows:
"Except for gender, all other sociodemographic characteristics were associated with the dietary contribution of ultra-processed foods. Both the crude and adjusted contribution decreased with age,

education, and income. The non-adjusted contribution was highest among Non-Hispanic Blacks (62.1 of total energy intake), followed by Non-Hispanic Whites (59.2 %), Mexican American (57.7 %), Other Hispanic (53.5 %) and Other (49.9 %). Adjusted estimates indicated similar consumption among Non-Hispanic Black and Non-Hispanic Whites but did not change substantially the consumption of the other ethnic groups. (Table 2).”

Table 2. I would have liked to have seen some estimate of precision, such as 99% confidence intervals for the values presented in this table – particularly the adjusted ones.

Authors: Thank you for highlighting this. We have now added the Confidence Intervals in this Table.

Line 302. I'd prefer to see the adjusted figures in the main text (rather than unadjusted), or at least in supplement data.

Authors. Thank you for your comment. After discussing the option of presenting the adjusted figures, we arrived at the conclusion that in this case it is more important to show observed food consumption rather than the estimated (adjusted) values. This is especially true if we consider that the sociodemographic characteristics of the population have not changed in this short period (results not shown).

Line 330-44. It would be nice to see some proposed explanations for the differences reported in this paragraph compared to the current results. Why might socio-economic gradients in UPF consumption be different in different countries? Why is UPF consumption higher in the USA even compared to Canada and the UK?

Authors: Thank you for your comment. Indeed, although there are slight differences in consumption of ultra-processed food among countries most of them present similar distribution of consumption across sociodemographic strata. Any speculation in regards to why UPF consumption is higher in the US than in other countries will be difficult to prove without understanding the determinants of consumption within each country. We consider this suggestion beyond the scope of this paper.

Line 345-56. Given that UPF are heterogeneous, I wondered whether the apparent difference between time trends in AHEI and UPF consumption may reflect differences in what UPF are consumed. This is a wider point that relates to the whole discussion – is it possible to choose more vs less ‘healthy’ UPF? If so, might this partly explain some of the e.g. relatively shallow socioeconomic gradients – more affluent people might be choosing a ‘healthier’ mix of UPF?

Authors: Thank you for highlighting this. We answered this question together with next point. Please see below.

Discussion. Throughout the discussion there is a tendency to compare the results to previous analyses of diet ‘quality’ in NHANES using a variety of other markers. This seems to assume that UPF consumption is just another marker of diet quality – is that the case? If so, I think a more robust discussion of what makes it better/worse than existing measures of diet quality would be helpful. If you are proposing UPF consumption as importantly different from other measures of diet quality, perhaps that could also be made more clear.

Authors: This point and the previous one are well taken. We have addressed them by rephrasing the paragraph

“While our study observed an increase in ultra-processed food consumption between 2007 and 2012, an investigation using the alternative index of diet quality (AHEI-2010) showed an increase in US adult diet quality between 1999 and 2012.[23] This stands in apparent contradiction if we take into account results from studies demonstrating that the dietary contribution of ultra-processed foods is inversely associated with the dietary content of protein, fiber, and most micronutrients and directly associated with carbohydrate, saturated fat, total sugar, added sugar and sodium contents.[14,24] However, it must be noted that more than 50% of the improvement in AHEI-2010 overall score was attributed to a reduction in trans fat intake,[25] which is not inconsistent with an increasing

consumption of ultra-processed foods, especially in view of the reduction in trans fat use in industrial products. The limits of ultra-processed food reformulation have been discussed elsewhere (Scrinis, 2016)

Line 393-404. You quote a newspaper article to claim that 'junk food' is not necessarily cheaper than 'real food'. There is abundant peer-reviewed evidence on this topic that would be preferable to cite.

Authors: Thanks for your comment. We added more articles reference on this paragraph and reformulate the phrase as follow:

"As previously highlighted in other studies, eating 'junk food' is not necessarily cheaper than eating 'real food'"

Line 410-12. Please provide evidence to support your claim that regulating marketing and promoting healthier food environments are the most effective way to change UPF consumption.

Authors: Thank you for highlight this. We already provide some evidence to support this idea on the paragraph when se state

"The influence of the school food environment, including school vending machines and school stores/ canteens/ snack bars, on children and adolescents' dietary intake is well known.[28]"

We have now added more references as follows:

Healthier eating should be encouraged among all sociodemographic strata especially by promoting healthier food environments (in schools among others) and regulating marketing[(Boyland,2016 ; Kubik,2010).

Line 426. Are region and urban/rural measured in NHANES? If so, could they not be included in your analyses?

Authors: Unfortunately this information is not available in NHANES.

Discussion. I'd like to see some consideration of the how e.g. 60% of energy from UPF should be interpreted from a clinical/public health point of view. Is this a 'healthy' amount? If not, what level should individuals and populations be aiming for? Similarly, it would be helpful to understand how important the statistically significant differences reported are. What difference in disease risk between e.g. 57.9% energy from UPF and 60.4% might be expected?

Authors: Thank you for highlighting this. In fact, we consider there is no healthy amount of ultra-processed food consumption, the less the better. In order to set a cut-off point for ultra-processed food consumption we would need to calculate a disease risk associated with each level of UPF consumption, which can only be explored in cohort studies.

#Reviewer 2

Line 108: omit 'and' and 'the'

Answer: Thank you for this correction. We sentence as follow:

NHANES is a continuous, nationally representative, cross-sectional survey of non-institutionalized, civilian US residents[15].

Authors use 'American White' 'American Black', but this is not consistent throughout the manuscript, and unconventional in general use. Suggest standardizing throughout using race/ethnicity terminology per NHANES. https://wwwn.cdc.gov/nchs/nhanes/2011-2012/demo_g.htm#RIDRETH1

Answer: Thank you for your comment. At first, we chose to use the terms adopted by Census also used in other articles (e.g. Powel et al, Prev Medicine, 2007). Following your advice, we have now adopted the NHANES terminology.

Assumptions: Categorizing foods into NOVA categories based on NHANES, which was not intended to facilitate such categorization, is complex. I might suggest, perhaps in the Discussion or Methods, listing some of the assumptions inherent in classifying, for example, 'cakes/cookies/pies' as 'ultra-processed', whereas it is reasonable possible that some homemade/bakery cakes may actually be in the 'processed' category. This might also be applicable to other UP foods. How did authors deal with this type of uncertainty? In how many categories/foods could this lead to classification bias?

Answer: Thank you for your suggestion. We have now added more detailed information in Methods (lines 194 to 212) :

"For all food items (Food codes) judged to be a handmade recipe, the classification was applied to the underlying ingredients (Standard Reference codes or SR codes) obtained from the United States Department of Agriculture (USDA) Food and Nutrient Database for Dietary Studies (FNDDS 4.0, 5.0 and 6.6) .For example, for cakes, cookies or pies underlying SR Codes were used unless underlying SR Codes or ingredients were unlikely to be used in home recipe (i.e. "Cellulose (Alpha-cellulose, powdered cellulose, and poly-cellulose)" or "Oil, industrial, soy (partially hydrogenated), multiuse for non-dairy butter flavor", or "Shortening, industrial, soybean (hydrogenated) and cottonseed", or "Whey, sweet, dried").

Other examples are "Salsa, red, cooked, not homemade" and "Salsa, red, cooked, homemade". Food Code "Salsa, red, cooked, not homemade" was classified as ultra-processed based on information from a similar product included in the Food code website ("Red Gold Salsa, Mild": tomato concentrate (water, tomato paste), diced tomatoes, jalapeno peppers, green chiles, yellow chiles, vinegar, salt, dried onion, dried garlic, cilantro, natural flavor). On the other hand, for Food Code "Salsa, red, cooked, homemade" it was each underlying SR Code (Salt, table; Garlic, raw; Onions, raw; Tomatoes, red, ripe, canned, packed in tomato juice; Peppers, hot chile, sun-dried; Water, tap, drinking; Vegetable oil, NFS) instead, that was classified according to NOVA, further explained in a previously published paper. [1]"

In lines 182-186, authors explain how they dealt with 'handmade' recipes, classifying underlying ingredients. In an average homemade meal that might fall into NOVA category 'processed' due to its incorporation of 'unprocessed' plus 'culinary ingredients', for example, would the approach taken by authors underestimate energy in the 'processed' category? For instance, would a homemade lasagna meal, instead of being categorized as 'processed', actually be classified as 'minimally processed' and 'culinary ingredients'? Please clarify.

Answer: Using your specific example, "Lasagna with meat" was assumed to be homemade and unfolded into underlying SR Codes (Cheese, mozzarella; part skim milk; Cheese, parmesan, grated; Cheese, ricotta, whole milk; Spices, garlic powder; Spices, oregano, dried; Salt, table; Onions, raw; Tomatoes, red, ripe, canned, packed in tomato juice; Tomato products, canned, paste, with salt added; Sugars, granulated; Macaroni, cooked, enriched; Ground beef or patty, cooked, NS as to percent lean (formerly NS as to regular, lean, or extra lean).

"Lasagna with meat, canned" on the other hand, was classified as ultra-processed food both based on the underlying SR Codes (Cheese, cheddar; Salt, table; Tomatoes, red, ripe, cooked; Water, tap, municipal; Syrups, corn, high-fructose; Modified food starch; Macaroni, dry, enriched; Beef, ground, 75% lean meat / 25% fat, crumbles, cooked, pan-browned) and ingredients of a similar product obtained from the Food code website (Stouffers Lasagna with Meat & Sauce: Sauce: tomato puree (water, tomato paste), water, cooked beef, dry curd cheese (cultured skim milk, enzymes), modified cornstarch, salt, bleached wheat flour, dried onions, sugar, spices, autolyzed yeast extract, dried garlic, soy sauce (water, soybeans, wheat, salt), dextrose, soybean oil. cooked pasta: water, semolina. Cheese: low-moisture part-skim mozzarella cheese (pasteurized part-skim milk, cheese cultures, salt, enzymes), part-skim mozzarella cheese and modified cornstarch (part-skim mozzarella cheese [pasteurized milk, salt, cultures, enzymes], modified cornstarch, nonfat milk, flavors), parmesan cheese (cultured milk, salt, enzymes), modified cornstarch).

Therefore, cases of uncertainty were solved using a conservative approach, opting for the lesser degree of processing or assuming a homemade recipe, which could have led to underestimation of ultra-processed food consumption.

We have now further explained this in the limitations as follows (line 438 -443):

“As NHANES was not specifically designed to classify food items according to degree of processing, misclassification errors may lead to under- or over-estimation of ultra-processed food consumption. Cases of classification uncertainty were solved using a conservative approach, opting for the lesser degree of processing or assuming a homemade recipe, which could have led to underestimation of ultra-processed food consumption.”

Results

Tables 1 and 2: would be helpful to add SD to means

Answer: Thank you for highlighting this missing data. We added SD to means on Table 1 and Confidence Intervals in Table 2

Table 2: footnotes are complex to interpret. Notably, * and ** seem to be inconsistently applied. Are they needed?

Answer. We appreciated your suggestion and made some modifications on the footnotes.

Line 258: should be ‘inversely’.

Answer: Thank you for pointing this out. We corrected this.

Line 259: do authors have an exact p-value (ie instead of listing <0.05)?

Answer: Thank you for your comment. In order to be more precise in this section we decided to add confidence intervals in Table 2.

Sentence on lines 264-266: Suggest omitting ‘both together’, replacing with something akin to: ‘However, differences in ultra-processed food consumption as a proportion of total energy between non-Hispanic Black and White groups disappeared, whereas consumption in both groups remained significantly higher compared to other race/ethnic groups’.

Answer: Following your suggestion, we have rephrased this entire section:

“Except for gender, all other sociodemographic characteristics were associated with the dietary contribution of ultra-processed foods. Both the crude and adjusted contribution decreased with age, education, and income. The non-adjusted contribution was highest among Non-Hispanic Blacks (62.1 % of total energy intake), then among Non-Hispanic Whites (59.2 %), Mexican American (57.7 %), Other Hispanic (53.5 %) and Other (49.9 %). Adjusted estimates indicated similar consumption among Non-Hispanic Black and Non-Hispanic White but did not change substantially the consumption of the other ethnic groups. (Table 2).”

Discussion

Might suggest comparing these NOVA findings with other NHANES studies in terms of: % energy from macronutrients, soft drinks, or other foods based on available data. This would help contextualize these findings with conventional methods of dietary assessment. For example, in lines 230-240 of Results, authors list distributions of energy intake according to NOVA- and in the Discussion, it might be helpful to see what other analyses not using NOVA have found. Authors can then compare, contrast, and speculate on reasons for the differences.

Answer: Although in Table 1 we do present subgroup energy contributions this was done to characterize UPF consumption pattern rather than to describe and focus on each individual subgroup energy contribution. As the focus of the paper was to describe the association between UPF consumption and sociodemographic characteristics we prefer to discuss these dietary aspects in another paper.

Line 319-320: “culinary recipes” is unclear, should this be “homemade foods/meals” or something similar? “turning them into an important eating pattern” is similarly unclear, I think because ‘them’ is improperly used.

Answer: Thank you for your comment. We have rephrased this sentence as follows:

“Small differences in consumption within the population reflect how ultra-processed foods have permeated and reached all social strata, modifying eating behaviors by displacing handmade meals, and turning ultra-processed food consumption into an important eating pattern.”

Line 321: ‘less’ should be ‘least’. Might suggest ‘individuals with a college education consumed the least...’ and not using the ‘exposed’ language, as NHANES measured consumption, not ‘exposure’ In other instances, authors should update to reflect ‘individuals with high school education’ and not ‘high school level individuals’, eg lines 326-327.

Answer. We appreciated your suggestion and corrected this paragraph as follows:

“Individuals with college education consumed the least ultra-processed foods whereas adolescents and American black and white ethnic groups were the highest consumers. Ultra-processed food consumption was inversely associated with both age and income levels and did not vary according to sex.”

As suggested, we have also substituted ‘high school level individuals’ by ‘individuals with high school education’ throughout the manuscript.

Lines 365-370: Sentence beginning “As a matter of fact...” is unclear. Perhaps something like, “Indeed, in recent years, patterns of food intake among children and adolescents have become less healthy”?

Answer. Thank you for your suggestion. We have modified this phrase as follows:

“The influence of the school food environment, including school vending machines and school stores/canteens/snack bars, on children and adolescents’ dietary intake is well known.[28] Indeed, in recent years, the healthiness of diets among children and adolescents are the ones to have declined the most.

Line 367-368: There is a subject-verb/grammar issue that makes this sentence unclear. Revise. Who/what had the greatest caloric increase? The children or the foods?

Also: what were the ‘ready-made Mexican dishes’?

Answer. Thank you for highlighting these points. We have rephrased this as follows:

“Snacks, pizzas, pastries, sweetened fruit juices and ready made Mexican dishes (mixed dishes with corn or flour tortillas and corn based dishes), most of them ultra-processed foods, were the ones to present the greatest caloric intake increase in the last 20 years among 2 to 6 year old American children.[29]”

Line 370: ‘fast foods’ instead of ‘fast food restaurants’

Answer. Following your suggestion, we have rephrased this sentence as follows:

“In addition, in the last decades among adolescents, a total of 17% of all calories came from fast foods.[24]”

Line 371: suggest beginning sentence with ‘Our findings are consistent with findings that...’

Answer. Thank you for your suggestion. We rephrased the sentence as follows:

“All these previous findings are consistent with our findings of higher ultra-processed food consumption among younger age groups and positive time trend in ultra-processed food intake among adolescents.”

Line 383: in an attempt to preserve cultural identity, yes, but also perhaps out of simple routine and not having been fully acculturated to the fast food culture yet.

Answer. Following your suggestion, we have completed this phrase as follows:

“Differences may also arise from the fact that households with foreign-born reference persons tend to cook more dinners at home than those with US-born reference person,[31] perhaps in an attempt to maintain their culinary traditions or out of simple routine and not having been fully acculturated to the fast food culture yet. However, it is known that throughout the process of acculturation (at least among Mexicans), migrants progressively incorporate negative eating habits, and with the passing of the generations end up adhering to the North American diet.[32]”

Line 390-391: suggest for clarity ‘may lead individuals to value nutrition over...’

Answer. As suggested we have rephrased this sentence as follows:

“Increased knowledge about nutrition and concern with health may lead individuals to value more nutrition over taste and convenience when choosing foods among more highly educated individuals.[26]”

Line 397: replace ‘breaks up’ with ‘counters’

Answer. As suggested we have replaced ‘breaks up’ with ‘counters’ as follows:

“On the other hand, the fact that only a small negative gradient was found between socioeconomic status and ultra-processed food consumption, counters the stereotype that lower income individuals are higher consumers of ultra-processed foods because of price.”

Line 405: omit ‘Even though’, begin sentence with “Socioeconomic disparities...” and begin new sentence with ‘Our study suggests’ (not shows)

Answer. Thank you for this suggestion. We have rephrased this section as follows:

“Socioeconomic disparities in health are thought to be partially caused by an SES gradient in nutrition.[30,33] Our study suggests that a gradient in ultra-processed food consumption is probably not the main or only underlying reason for these health inequities.”

Line 408: suggest replacing ‘gaps’ with ‘inequities’

Answer. Thank you, we have followed your suggestion as shown above.

Also: revise to read ‘high overall’

Answer. Following your advice, we have rephrased this sentence as follows:

“Indeed, ultra-processed food consumption was high overall, despite being slightly higher among less educated, younger, lower income and American black and white strata.”

Line 409: replace ‘superior’ with ‘higher’

Answer. Thank you, we have followed your advice as shown above.

Line 410: replace ‘fomented’ with ‘encouraged’

Answer. We agreed and have rephrased this sentence as follows:

“Healthier eating should be encouraged among all sociodemographic strata especially by promoting healthier food environments (in schools among others) and regulating marketing.”

VERSION 2 – REVIEW

REVIEWER	Jean Adams Centre for Diet & Activity Research, University of Cambridge, UK
REVIEW RETURNED	12-Jan-2018
GENERAL COMMENTS	Thanks for responding to my previous comments.