What kind of evidence would be most persuasive to reduce global meat consumption?

Lee Moore, Sonal Thakkar, Peter Trenkwalder
W241 Summer 2019 Section 3 Final Project

1 Background and Research Question

There is growing evidence that meat consumption leads to worse health (WHO, 2015) and has important negative effects on the environment (UN IPCC, 2019). However reducing global meat consumption is a major challenge. Firstly, because it is nearly impossible to conduct long-term randomized controlled experiments on meat consumption, the causal impact of meat on long-term health can be easily disputed. Secondly, because population studies in nutritional science often conflict with each other, and because there are valid reasons to believe population level effects may not be relevant to individuals due to considerable heterogeneity in individual genetic and environmental propensity towards nutritionally-driven morbidity, it is easy to dismiss these findings as personally irrelevant. Finally, regarding the impact on the environment, given the existing skepticism on climate change itself, there may be limited impact of highlighting the negative impact of livestock on the environment in order to drive a change in meat consumption behavior.

Our research study is based on the premise that, with a goal to reduce global meat consumption and with limited resources to develop evidence to support this goal, it is important to identify what kind of evidence would provide the most impact. Our research question therefore asks: *What kind of evidence would be most impactful to reduce global intended meat consumption?* The three types of evidence evaluated in this study are:

A. irrefutable population health evidence of the negative impact of meat
B. personalized nutritional evidence based on genetic sequencing and advanced analytics
C. irrefutable climate change evidence of the negative impact of meat

Our hypothesis is that the existence of all three types of evidence would have a positive impact on individuals' willingness to attempt to decrease their level of meat consumption. In addition, we are interested in understanding whether a specific evidence type stands out as having the strongest positive impact on individual's willingness to decrease their meat consumption. Previous research by Celis-Morales *et al* (2017) has found that personalized nutritional advice (similar to Evidence Type B) was more effective in improving dietary behaviours when compared with conventional 'one size fits all' population-based advice (similar to Evidence Type A). It will be interesting to understand how these two types of evidence 'rank' in our study in terms of importance in decreasing intended meat consumption, and, relative to these, where environmental evidence (Evidence Type C) would fall.

2 Methodology

2.1 Potential Outcomes

The outcome of interest in our study is a change in intended level of meat consumption. The five ordinal levels tested were the following:

1. Regular Meat Eater (2 or more meals a day with meat included)
2. Moderate Meat Eater (1 meal a day with meat included)
3. Low Meat Eater (less than 1 meal a day with meat, so some days I don't eat meat)
4. Vegetarian (No meat, but still some animal-related products like milk or eggs)
5. Vegan (No meat or animal-related products)

Study subjects were asked to state the category that best describes their current level of meat consumption and were then asked to state their preferred level of meat consumption, using these same five levels, following their randomized treatment. We used these 5 descriptive levels of meat consumption because this was easier for our study subjects to understand and answer; rather than specific continuous quantities of meat, for instance, a change in intended total average grams of meat consumed per day.

The potential outcomes were evaluated in two ways:

- Using proportional odds logistic regression, considering specifically which level of meat consumption a study subject stated at baseline and which level they ended in following treatment
- Using linear regression, a binary outcome on whether the new intended level of meat consumption following treatment was less than (tagged as 1), or greater than or equal to (tagged as 0) their baseline level of meat consumption

2.2 Population

Our population of interest is the general population of those who consume any meat or animal-related products. Our survey subjects came from Amazon's Mechanical Turk (Mturk) platform where we offered a payment of 40 cents per completed survey. Several baseline covariates were collected in order to understand the demographic characteristics of our study subjects. These were age, gender, income, education and current intention to change meat consumption - see the Appendix for full details on the questions and answer categories included in the baseline survey.

The only exclusion criteria applied was for those who stated "Vegan" as their baseline meat consumption level. Since the purpose of the experiment was to measure the reduction in meat consumption based on the treatment panels, vegans were excluded due to lack of meaningful potential outcomes based on our study objectives. Specifically, their survey ended after completing the baseline survey details and they were not included in the randomization to treatment panels.

Data analysis was restricted to those who were defined to be compliant to treatment, which will be discussed further in section 2.3.

2.3 Treatment

Following collection of baseline meat consumption level as well as other baseline covariates, study subjects were block randomized based on their stated baseline meat consumption level to see two of four panels sequentially. Each panel contained a brief passage and some images meant to support understanding of the passage's content. In the case of the placebo panel (hereon referred to as panel X), only images were shown. The three treatments evaluated in our study are aligned to those described in section 1 and will hereon be referred to as panels A, B and C:

- Panel A. **Population health evidence** irrefutably confirming the negative impact of meat on population health through increased cancer, heart disease and general acceleration of the aging process
- Panel B. **Personalized nutritional evidence** outlining a situation where advances in genetic sequencing and advanced analytics provides individual tailored nutritional advice that states that if the study subject eliminates meat, s/he will reduce their chance of early death from cancer and heart disease
- Panel C. **Environmental health evidence** irrefutably confirming the negative impact of livestock on the environment

Following review of the panel, subjects were then asked to complete three tasks on the assumption that the evidence presented was true:

1. answer a comprehension check question to ensure they reviewed the passage/images,
2. provide free text for sentiment analysis, and finally
3. select their preferred level of meat consumption.

These four panels and associated questions have been replicated in the Appendix. The purpose of requiring study subjects to assume the evidence was true was to remove variability which would exist due to individual's differing degrees of skepticism towards each evidence type and focus attention instead on the nature of the evidence itself. This is discussed further in section 4.

A 2-staged pilot survey was used for the purposes of refining our survey design. The survey text, questions and panels was first piloted to friends and family to get feedback and then opened to a group of 40 MTurk users. All text and images used as part of our final study were therefore subjected to various iterations and improvements to ensure clarity and simplicity.

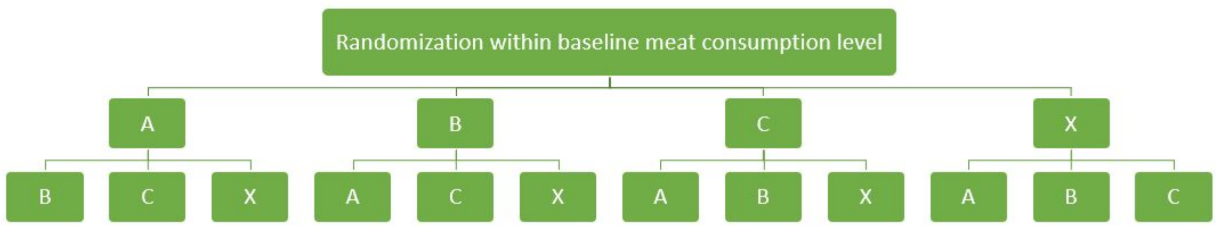
The first question following each panel was used as our compliance check. These were designed to be sufficiently basic such that a brief review of the text and/or images would be sufficient to answer this correctly. We required that an individual answer this question correctly for both of their randomised panels in order to be considered compliant and appropriate for evaluation within our analysis. This was to minimise the probability that those who did not review the panels would guess the correct answers and be incorrectly classified as a complier. Section 3.1.1 below provides an assessment of the compliance rate in our study.

2.4 Randomization Process

Another important learning from the pilot study was that we could expect a high level of attrition (80% observed in our pilot) unless we combined our baseline survey with the randomized panels. Our original survey was designed using html within Mturk as two separate surveys, so that we could perform a block randomization procedure ourselves based on the information gathered within the baseline survey. While the pilot's first (baseline) survey was completed by 40 survey subjects, the second (randomized) survey administered later saw just 8 MTurk users from the original 40 respond. We therefore decided to switch to Qualtrics as a survey platform since Qualtrics allows for randomization within the survey, and while we lose some control over the randomization procedure, we benefit from limited attrition due to survey design.

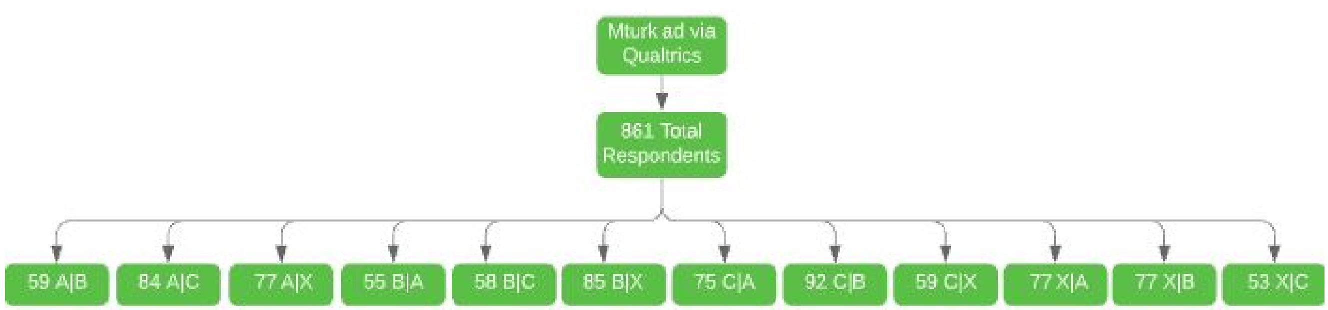
Qualtrics was therefore used for the purposes of both data collection and randomization. Following collection of baseline meat consumption level as well as other baseline covariates, those who selected "Vegan" were excluded from the randomization while the remaining study subjects were block randomized based on their stated category to receive one of four interventions (either to read one of three evidence panels, A, B, or C or the placebo panel X). Following this first randomization, study subjects were then randomized again to receive one of the remaining three panels they had not already seen in their first treatment, followed by a similar set of questions. This flow is represented in Figure 1.

Figure 1. Representation of randomization possibilities for the first and second randomizations



Qualtrics randomization procedure allows for a nearly complete randomization when the "Evenly present elements" option was selected. Essentially this meant that simple randomization was applied initially, with subsequent treatments allocated in such a way to ensure that total counts for panels A, B, C and X within each baseline meat consumption category were balanced. Once equilibrium of the four panel counts was achieved, simple randomization would again be applied to the subsequent randomizations, and so on. The final aggregate counts for the sequence of randomization possibilities are shown in Figure 2 below.

Figure 2. Flow Diagram of aggregate treatment allocation (ignoring blocking by baseline meat consumption)



Ideally, the randomization would have occurred such that an equal number of study subjects would be allocated to each of the 12 possible permutations of treatment assignments, however Figure 2 demonstrates that this was not the case. Given that every study participant was randomized to see two panels, this procedure in Qualtrics ensured this balance occurred over the combined counts of panels across the two randomizations. This can be seen nicely in the example of the "Low Meat Eater"

Category which had at total of 194 survey subjects and therefore 388 (194 x 2) panels evaluated. While those randomized to A, B, C, and X were not perfectly balanced in the first randomization (counts were 54, 47, 52, and 41 respectively) nor in the second randomization (counts were 43, 50, 45, and 56 respectively), when these counts were combined, we have exactly 97 counts in each of the 4 panels. While none of the other three meat consumption categories have such precise counts, it can be assumed that they would have met a similar equilibrium if we had additional study participants.

Table 1. Total counts of treatment allocation for all four baseline levels of meat consumption

Regular Meat Eaters

	Second Randomization				
First Randomization	A	B	C	X	Total
A		13	20	23	56
B	8		15	23	46
C	20	29		10	59
X	25	14	12		51
Total	53	56	47	56	212

	A	B	C	X	Total
Combined Counts	109	102	106	107	424

Moderate Meat Eaters

	Second Randomization				
First Randomization	A	B	C	X	Total
A		25	36	21	82
B	27		23	39	89
C	35	30		28	93
X	30	38	21		89
Total	92	93	80	88	353

	A	B	C	X	Total
Combined Counts	174	182	173	177	706

Low Meat Eaters

	Second Randomization				
First Randomization	A	B	C	X	Total
A		12	20	22	54
B	12		14	21	47
C	17	22		13	52
X	14	16	11		41
Total	43	50	45	56	194

	A	B	C	X	Total
Combined Counts	97	97	97	97	388

Vegetarians

	Second Randomization				
First Randomization	A	B	C	X	Total
A		9	8	11	28
B	8		6	2	16
C	3	11		8	22
X	8	9	9		26
Total	19	29	23	21	92

	A	B	C	X	Total
Combined Counts	47	45	45	47	184

The consequence of the lack of balance of treatments administered for each meat consumption level is that any analysis or summary statistic which attempts to pool data from all four baseline meat consumption categories together would need to account for the different proportions of treatment panels administered for each baseline meat eating category to ensure bias is not introduced.

2.5 Statistical Calculations

Prior to our study, we calculated that based on the budget available for our study, approximately 800 Mechanical Turk survey subjects are feasible to recruit, and with an assumed compliance rate of 75%, there would be approximately 600 survey subjects for evaluation. We are interested in the power calculation associated with the individual baseline meat categories, since the randomization was blocked on this attribute and we would like to detect an effect within each subgroup. Assuming (simplistically) an equal number of study subjects per baseline category (n=150 survey subjects per baseline meat consumption category) and that each survey subject provides an evaluation of 2 panels (n=300 panels evaluated per baseline meat consumption category), this means we would have approximately 75 responses per panel (300 responses / 4 panels).

Power calculations were based on the hypothesis that each of the three treatment panels will have a significant effect over placebo, rather than on our secondary interest in potentially observing a larger treatment effect for specific treatment panels over others. Assuming we would see an effect size of 50% for each treatment panel over the placebo panel, using a two-sample t test power calculation with n=75 in each group, our experiment anticipates generating a power of 0.860 with alpha = 0.05 which is above the conventional power of 0.80.

3 Analysis

3.1 Descriptive Statistics / EDA

3.1.1 Recruited Population and Compliers

Figure 3 shows the flow of survey subjects according to baseline meat consumption, providing the total counts of those who completed the survey as well as total counts for those defined as compliers who will be used for our statistical analysis. Total counts of first and second randomization assignments are provided in Table 1 above (within section 2.4) and the total counts of compliers for each randomization assignment are provided in Table 2 below.

Figure 3. Flow Diagram of survey subjects by baseline meat consumption

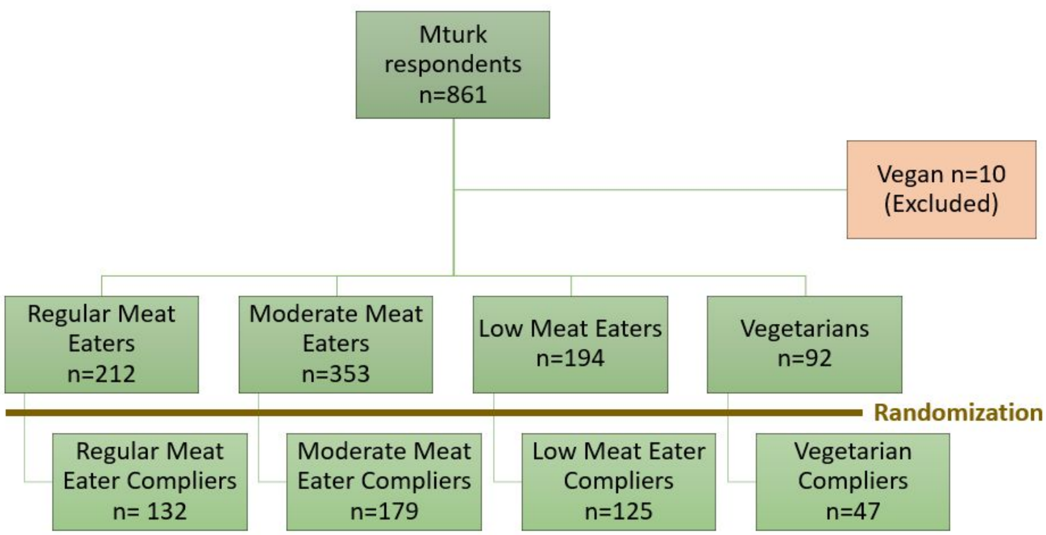


Table 2. Total counts of compliers within each treatment allocation permutation for all four baseline levels of meat consumption

Regular Meat Eaters

First Randomization \ Second Randomization	A	B	C	X	Total
A		10	13	16	39
B	3		7	13	23
C	14	20		6	40
X	12	9	9		30
Total	29	39	29	35	132

Moderate Meat Eaters

First Randomization \ Second Randomization	A	B	C	X	Total
A		9	17	11	37
B	11		7	22	40
C	20	11		23	54
X	11	20	17		48
Total	42	40	41	56	179

Low Meat Eaters

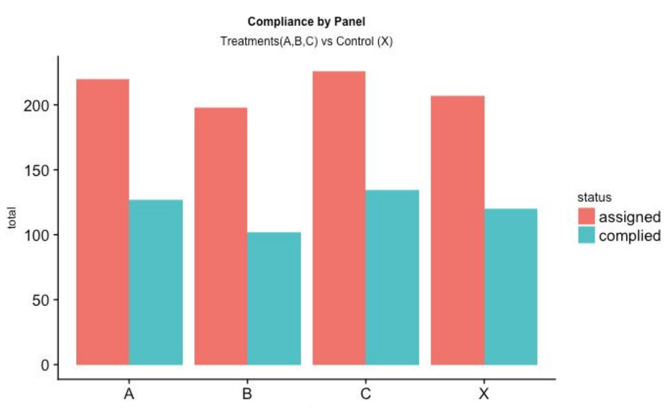
First Randomization \ Second Randomization	A	B	C	X	Total
A		7	11	16	34
B	10		8	13	31
C	10	10		10	30
X	11	11	8		30
Total	31	28	27	39	125

Vegetarians

First Randomization \ Second Randomization	A	B	C	X	Total
A		4	5	8	17
B	3		4	1	8
C	1	3		6	10
X	4	0	8		12
Total	8	7	17	15	47

Compliance, while low, seemed reasonably consistent and comparable across panels as seen in Figure 4 below. However a 2-sample test for equality of proportions between the treatment (combined panels A/B/C) and placebo (panel X) reveals a statistically significant ($p=0.01494$) difference in the proportion of compliers in the placebo group (61.9%) compared to the treatment groups (55.0%). Given that our placebo group only had to review and correctly identify images, rather than read a text and answer a reading comprehension question correctly, it is not surprising that compliance rates overall for placebo are higher and this is a limitation of our study discussed further in section 4.

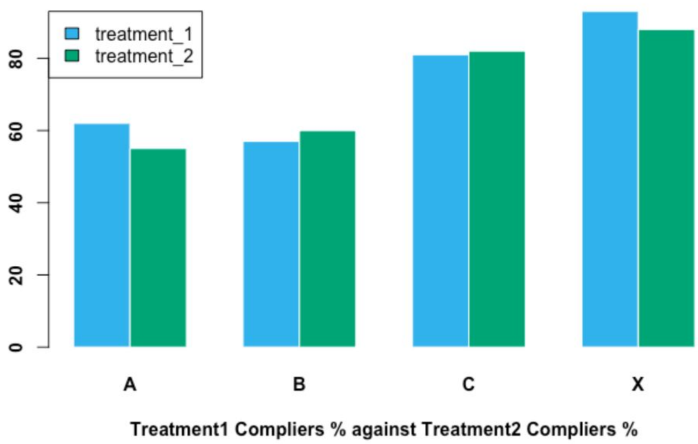
Figure 4. Assignment and Complier counts per panel for 1st Randomization



While this may suggest that our ‘apples-to-apples’ comparison for our complier average causal effect calculations is at risk, it is important to recall that each survey subject was asked to complete two randomized panels, which means that someone who was assigned the placebo (panel X) was also assigned to one of the three treatment panels (panel A, B, or C). As a consequence, a survey subject who did not read the text associated with their treatment panel but did review the images would have answered the placebo compliance check question correctly, but would have still been categorized overall as a non-complier if they did not answer the treatment (panel A, B, or C) compliance check question correctly. While the probability of guessing the correct answer to a treatment panel is higher for those who were randomized only to see one treatment panel and the placebo panel (rather than two treatment panels), we believe that an apples-to-apples comparison is still maintained amongst compliers given this study design.

We also evaluated the compliance rate between the first and second randomizations as shown in Figure 5 below. Whilst compliance levels decreased for panels A and X, they increased for panels B and C. Consequently, there was no clear pattern of reduced compliance between the first and second randomizations. All subsequent analyses focuses on compliers only.

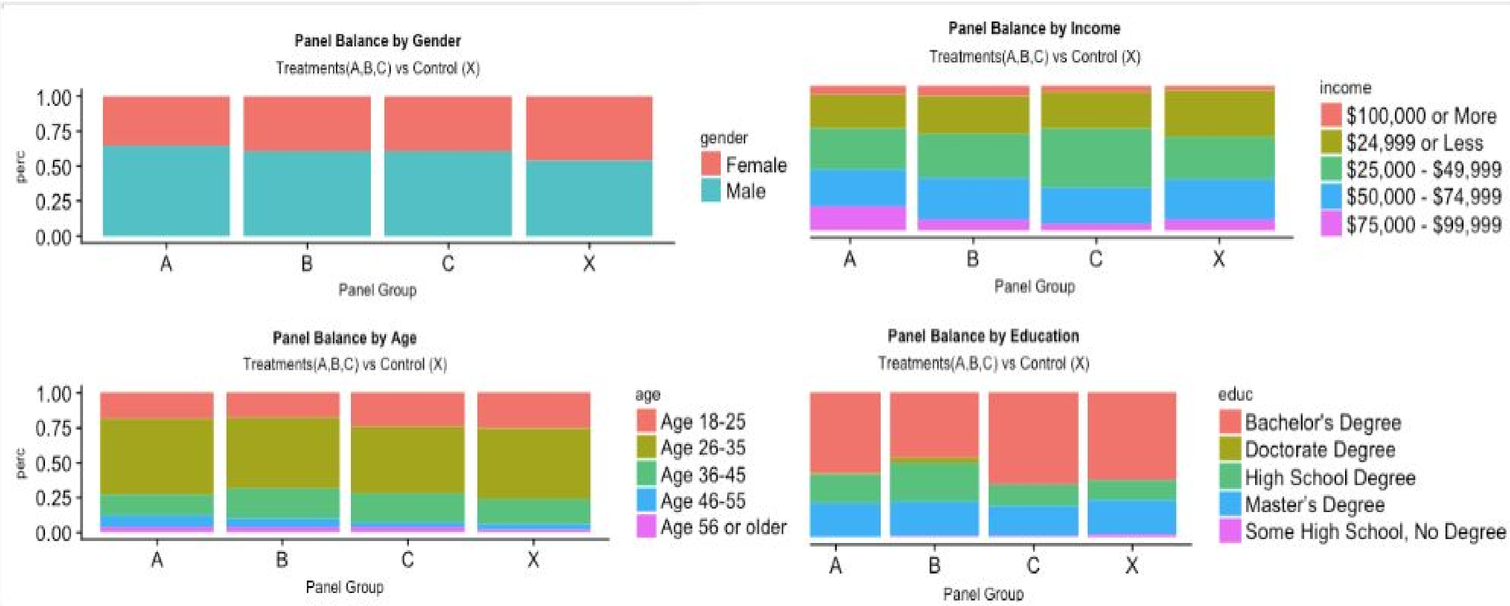
Figure 5. Proportion of compliance to each panel for first treatment assignment versus second treatment assignment



3.1.2 Covariate Balance

Our covariates appear well balanced across panels (see Figure 6 below) suggesting an effective randomization and the ability to make apples-to-apples comparisons across various demographic groups.

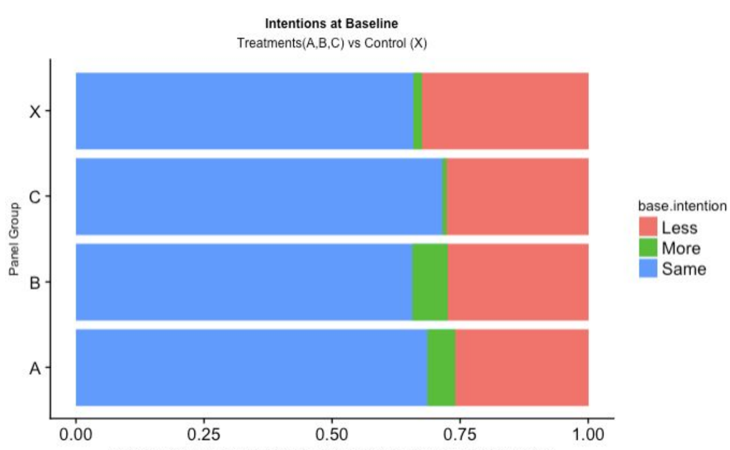
Figure 6. Covariate distribution by 1st Panel for Compliers



3.1.3 Baseline Intentions

From Figure 7, we can see that a substantial proportion of compliers already intend to lower meat consumption (eat *Less*) without any intervention. To ensure we do not overstate treatment effects, we need to account for baseline intentions and the relative differences between panels.

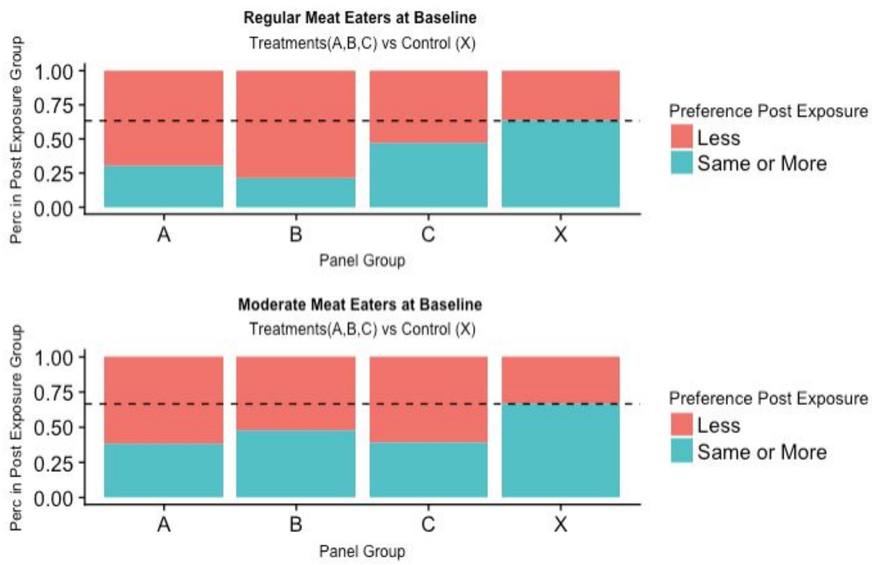
Figure 7. Baseline intention to change meat consumption for compliers by 1st randomized treatment



3.1.4 Changes Post Exposure to Treatments and Control

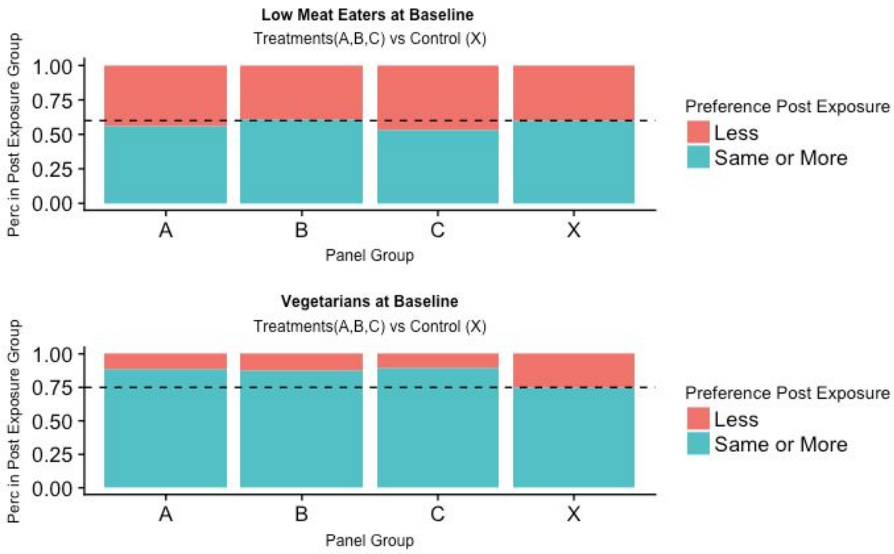
A noticeable difference in preference exists between the first randomized treatment panels A, B, C and control (panel X) after intervention among those who were Regular and Moderate meat eaters at baseline as shown in Figure 8. This seems suggestive that exposures to panels A, B, or C may have more of an effect than control (panel X) at influencing these types of meat eater's desire to lower their consumption

Figure 8. Proportion of Regular/Moderate Meat Eaters who would intend to reduce their meat consumption post-first randomized treatment, by Panel (Compliers only)



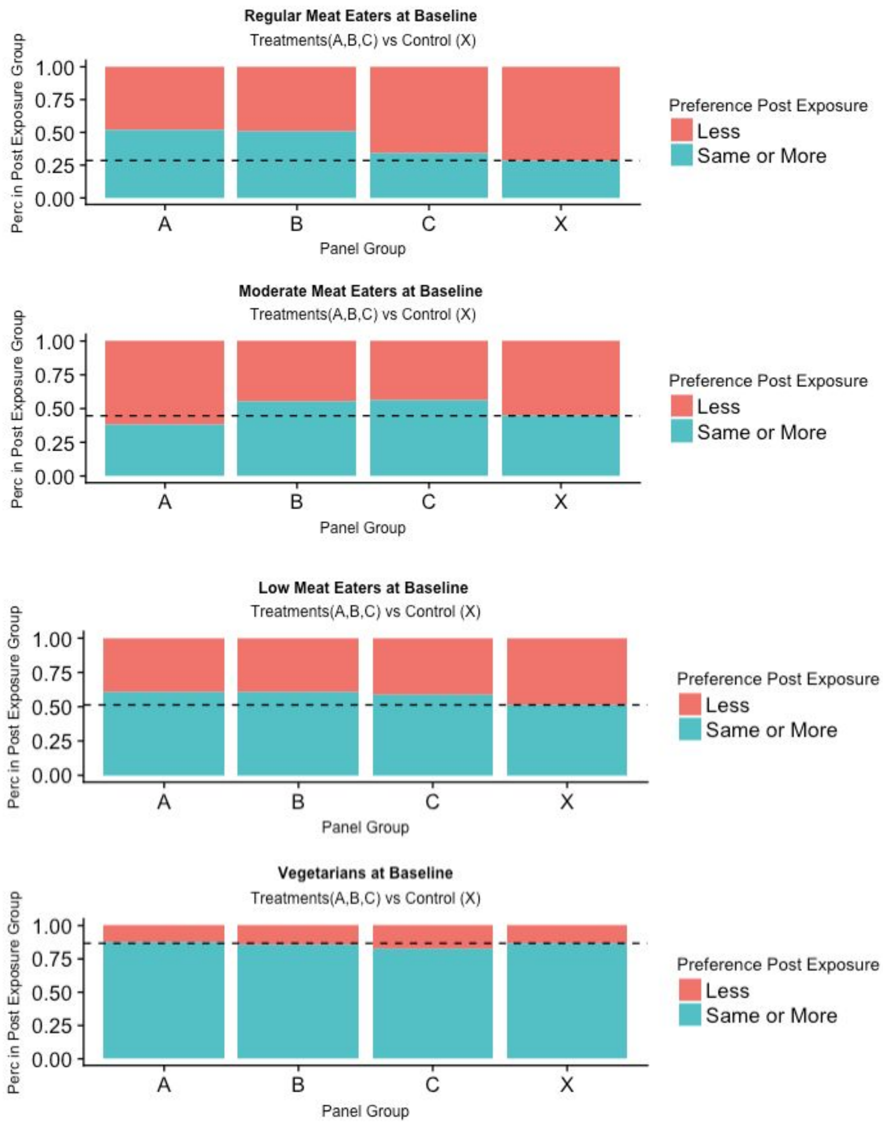
The noted effect above does not appear evident among Low to Moderate Meat Eaters as shown in Figure 9 when comparing treatment groups to control. This is not too surprising given that those who are already low meat eaters or vegetarians may be more cognizant of the downsides to meat consumption and, therefore, may base their rationale on similar intuition as what was presented in the evidence panels.

Figure 9. Proportion of Low Meat Eaters/Vegetarians who would intend to reduce their meat consumption post-first randomized treatment, by Panel (Compliers only)



As shown in Figure 10 below, among baseline *regular* and *moderate* meat eaters, differences between control and treatment panels for the second randomization are less evident than what we observed when these panels were administered in the first randomization. What stands out here is the major change in the proportion of those who would intend to eat less meat post exposure in our second randomization control group (panel X) compared to the proportion observed in our first randomization control group (panel X). This is counter-intuitive because we would not expect control behavior to change. However, it's possible that the first randomization panels may be priming compliers to respond more favorably in the second randomized panels so we will check for this interaction as part of the statistical analysis. Among baseline *low* meat eaters and *vegetarians*, the proportion of those who would intend to eat less meat post exposure in our second randomization panels seem consistent with the proportion observed in our first randomization panels.

Figure 10. Proportion of each baseline meat consumption group who would intend to reduce their meat consumption post-second randomized treatment, by Panel (Compliers only)



3.1.5 Sentiment Scores Among the Treated

Those who were treated and wanted to eat less compared to those who were treated and wanted to eat the same or more yielded a t-statistic of 2.32 with a p-value of .02 after testing for a difference in negative sentiment between the groups. A statistically significant greater degree of negative sentiment toward panel content exists in the group who would eat less.

Figure 11. Sentiment analysis of compliers who intend to eat less v. same/more post-treatment

Treated who Would Eat Less

Sentiment	Negative	Neutral	Positive
Score	.15	.68	.17

Treated who Would Eat the Same or More

Sentiment	Negative	Neutral	Positive
Score	.11	.66	.23

3.2 Statistical Analysis

While our outcome data is ordinal which allows us to generate a model that provides insight into the probabilities of selecting one type of meat consumption category over another for a given panel and baseline, our primary concern is whether or not our treatment panels have any influence on lowering meat consumption among compliers. As such, we initially fit our data against a linear model to test for this. Afterward, we used proportional odds logistic regression to better understand the chances of being a particular type of meat eater given a treatment compared to control. As noted previously, all analyses are based on compliers only rather than the overall intent-to-treat population.

3.2.1 Modeling with Linear Regression

Our linear regression fit our treatment variables on a binary outcome after controlling for a binarized baseline intention. Any complier who at baseline intended to eat less meat was flagged as 1 while those who intended to eat the same or more meat were flagged as 0. Similarly, any complier whose outcome indicated they would prefer less meat than their baseline level was flagged as 1 or 0 otherwise. Our first model can be viewed as:

$$Y = \beta_0 + \beta_{Treatment} + \beta_{BaselineIntent} + e$$

$Y =$ *expected outcome of treatment group above or below control with respect to eating less meat*

$\beta_0 =$ *expected outcome of control (i.e. panel X) with respect to eating less meat*

Table 3. 1st Treatment Linear Regression on intent to reduce meat consumption post-treatment

1st Treatment	Estimate	Std. Error	t-value	Pr(>|t|)
Placebo X (intercept)	0.33	0.045	7.28	1.3e-12*
A	0.18	0.062	2.89	0.003*
B	0.16	0.066	2.45	0.014*
C	0.16	0.061	2.72	0.006*
Baseline Intention	0.05	0.051	1.12	0.26

Table 3 above provides the results of our overall complier regression based on the first randomized treatment. On average, we expect panel A to generate an effect on lowering intended meat consumption 18 percentage points above the effect of control. For panels B and C, we expect the effect on lowering intended meat consumption to be 16 percentage points above control. So, for example, on average 51 percent of people initially exposed to panel A report to want to lower meat consumption while only 33 percent of people initially exposed to control report to want to lower meat consumption. We can see that there is very little chance these estimates are derived from sampling variability since all p-values are significant at less than a .05 level. This model has adjusted for baseline intention to reduce meat consumption which was not found to be statistically significant with a p-value > 0.05. Generally, we do see an average treatment effect among our compliers.

Baseline Categories Driving Observed Effects

What we do not see in the table above is which baseline category of meat eaters are driving this effect. Our EDA showed that the differences from control seemed greater among Regular and Moderate meat eaters. We refit the linear model above for each of the baseline categories of meat eaters and found the Low meat eaters and Vegetarians to have no statistically significant coefficients, meaning the Regular and Moderate meat eaters are driving this effect. These pre-defined subgroup analyses are presented in Table 4 below.

Table 4. 1st Treatment Linear Regression on intent to reduce meat consumption post-treatment by subgroups of baseline meat consumption

	Dependent variable:			
	Regular (1)	Moderate (2)	Low (3)	Vegetarian (4)
treatment_1A	0.360***	0.299***	0.040	-0.132
	(0.116)	(0.109)	(0.126)	(0.163)
treatment_1B	0.443***	0.173	0.012	-0.122
	(0.125)	(0.108)	(0.127)	(0.195)
treatment_1C	0.193	0.264***	0.072	-0.149
	(0.119)	(0.098)	(0.128)	(0.182)
base.intention.binary	0.167	0.120	0.174*	0.007
	(0.125)	(0.091)	(0.095)	(0.116)
Constant	0.311***	0.313***	0.325***	0.245
	(0.088)	(0.072)	(0.095)	(0.161)
Note:	*p<0.1; **p<0.05; ***p<0.01			

Differences between Treatment Panel Effects

We re-leveled our intercept to represent one of the three treatment panels (A,B or C) to test if some treatment panels had greater effects on lowering meat consumption than others. The below results in Table 5 show that none of the treatment panels had a statistically significant different effect from the others on changing consumption behavior.

Table 5. Regression from Table 3 reproduced with intercept set at Panel A, B, and C respectively

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	0.527559	0.044652	11.8149	< 2.2e-16	***
treatment_1X	-0.177559	0.062623	-2.8354	0.004771	**
treatment_1B	-0.017755	0.067027	-0.2649	0.791204	
treatment_1C	-0.012634	0.062337	-0.2027	0.839482	

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	0.5098039	0.0499879	10.1985	< 2e-16	***
treatment_1A	0.0177551	0.0670268	0.2649	0.79120	
treatment_1X	-0.1598039	0.0665329	-2.4019	0.01669	*
treatment_1C	0.0051215	0.0662642	0.0773	0.93843	

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	0.5149254	0.0434988	11.8377	< 2.2e-16	***
treatment_1B	-0.0051215	0.0662642	-0.0773	0.938426	
treatment_1A	0.0126337	0.0623373	0.2027	0.839482	
treatment_1X	-0.1649254	0.0618059	-2.6684	0.007879	**

Covariate and Heterogeneity Analysis

To test whether we could improve the precision of our complier average causal effect (CACE) we modeled a second linear regression to control for our demographic covariates as:

$$Y = \beta_0 + \beta_{Treatment} + \beta_{BaselineIntent} + \beta_{Age} + \beta_{Gender} + \beta_{Education} + \beta_{Income} + e$$

Covariates were included in both their original categorical form (see Appendix for categories) as well as binarized based on appropriate splits following assessment of histograms. However, our covariate analysis did not yield any statistically significant results that would help generate more precise estimates of the CACE.

We also tested for heterogeneous treatment effects among our demographic covariates with the below model, but again found no statistically significant results that could provide insight into whether the effect was greater for a particular subset of our demographic categories.

$$Y = \beta_0 + \beta_{Treatment} * \beta_{Age} + \beta_{Treatment} * \beta_{Gender} + \beta_{Treatment} * \beta_{Education} + \beta_{Treatment} * \beta_{Income} + \beta_{BaselineIntent} + e$$

Interactions between Treatment 1 and Treatment 2

Because our survey design consisted of delivering an initial panel followed by a second panel, it is plausible that respondents outcomes to the second panel were primed by exposure to the first panel. We explored the interaction between the sequence of treatments to delineate if a particular sequence of panel content was more or less effective at lowering meat consumption than others. We did this via an analysis of variance test between the two models below:

$$Y = \beta_0 + \beta_{Treatment2} + \beta_{BaselineIntent} + e \qquad \textbf{vs} \qquad Y = \beta_0 + \beta_{Treatment1} * \beta_{Treatment2} + \beta_{BaselineIntent} + e$$

The results of our analysis of variance yielded an F-score of .4418 and a p-value of .89 and therefore we fail to reject the null hypothesis that the two models are the same. This suggests the simpler model would be preferred to the more complex model, as the latter does not provide a better overall model fit (i.e. the inclusion of the interaction term with treatment 1 does not help us explain more of the variation in the outcome). What this means is that there is no significant difference between any given sequence so we can evaluate treatment 2 independently of treatment 1 as shown in Table 6 below.

Table 6. 2nd Treatment Linear Regression on intent to reduce meat consumption post-treatment

2nd Treatment	Estimate	Std. Error	t-value	Pr(>|t|)
Placebo X (intercept)	0.40	0.044	9.07	2e-16*
A	0.14	0.063	2.28	0.022*
B	0.11	0.062	1.78	0.074
C	0.015	0.061	0.24	0.80
Baseline Intention	-0.04	0.05	-.96	0.33

Panel A, when shown as the second treatment, has the only significant impact with respect to lowering reported meat consumption. This was surprising at first because panels B and C, when showed as the first treatments, were significant and there was no identified interactions between the sequence of treatments. Referring back to our EDA when considering how preferences post-exposure differed between panels based on whether they were shown as the first or second randomized treatment, we found that preferences toward less consumption appeared to change most among Regular and Moderate meat eaters in the control. Given that those who received the control panel X as the second treatment were initially exposed to a treatment panel (panel A, B, or C) in the first randomization, these subjects might be responding based on what they believe to be appropriate behavior. In other words, a demand effect could be adding bias into our experiment which would help explain why we found no significant interactions with the first randomization and, yet, fewer panels in the second randomized treatment had a statistically significant impact on lowering meat consumption.

3.2.2 Modeling with Proportional Odds Logistic Regression

Our original outcomes (vegan, vegetarian, low, moderate, and regular meat eaters) are ordinal in that there is a known ranking among these in terms of meat consumption, but the differences between each are not known and are likely inconsistent. For instance, it may be much easier for a moderate meat eater to convert to a low meat eater than for a low meat eater to convert to a vegetarian. These types of outcomes can be modeled via proportional odds logistic regression (polr) as shown in Table 7 so we can understand the probability of being in a particular category given a treatment, baseline consumption, and baseline intention.

Table 7. 1st Treatment Proportional Odds Logistic Regression on intent to change meat consumption level post-treatment

```
Call:
polr(formula = factor(outcome_1, levels = c("Vegan", "Vegetarian",
    "Low", "Moderate", "Regular")) ~ treatment_1 + base.intention.binary +
    base.consumption, data = dt %>% filter(both_correct == 1),
    Hess = TRUE)

Coefficients:
                          Value Std. Error t value
treatment_1C            -0.6484     0.2316  -2.799
treatment_1B            -0.8545     0.2480  -3.445
treatment_1A            -0.8856     0.2394  -3.699
base.intention.binary   -0.5330     0.1939  -2.748
base.consumptionLow      0.9803     0.3187   3.076
base.consumptionModerate 2.3872     0.3270   7.299
base.consumptionRegular  3.4979     0.3537   9.889

Intercepts:
                 Value   Std. Error t value
Vegan|Vegetarian -2.0279  0.3432    -5.9086
Vegetarian|Low    0.0416  0.3364     0.1238
Low|Moderate      2.0153  0.3506     5.7486
Moderate|Regular  3.5842  0.3645     9.8327

Residual Deviance: 1254.067
AIC: 1276.067
```

The coefficient outputs from our polr model above are in the form of log odds which are difficult to interpret as is. We can exponentiate these into odds form and then convert the odds into probabilities which provides us with the probability that someone who identifies as a particular meat eater at baseline, and who has some baseline intention to change their meat consumption, would be influenced by a given panel to prefer to be a different level of meat eater.

For instance, what we are specifically interested in is the ability of our treatments to influence individuals who identify as regular meat eaters with no intention of lowering their meat consumption to want to lower their meat consumption after exposure to our treatment panels. We can gain insight into this situation using R's predict function with our polr model fit, passing in the parameters of interest.

````
```{r}
predict(polr_fit,newdata = data.frame(treatment_1="A",base.consumption = "Regular",base.intention.binary = 0),type="p")
predict(polr_fit,newdata = data.frame(treatment_1="B",base.consumption = "Regular",base.intention.binary = 0),type="p")
predict(polr_fit,newdata = data.frame(treatment_1="C",base.consumption = "Regular",base.intention.binary = 0),type="p")
predict(polr_fit,newdata = data.frame(treatment_1="X",base.consumption = "Regular",base.intention.binary = 0),type="p")
```
````

Each call to predict will generate the probability of being in one of the 5 consumption categories given the parameters passed. From the calls above we generated the following table of probabilities as shown in Table 8.

Table 8. Post Exposure Probabilities for Regular Meat Eaters with No Intention to Lower Consumption at Baseline by Panel for first randomization

post_preference <chr>	panel_a <dbl>	panel_b <dbl>	panel_c <dbl>	control <dbl>
Vegan	0.009564071	0.009273794	0.007559281	0.003967108
Vegetarian	0.061489238	0.059753059	0.049340535	0.026616436
Low	0.283979774	0.278914186	0.245847887	0.154451583
Moderate	0.370467639	0.371320092	0.373078017	0.336527031
Regular	0.274499278	0.280738869	0.324174281	0.478437842

We can read the above as: for regular meat eaters with no intention of lowering meat consumption at baseline who were exposed to panel A, there is a 27% chance they would intend to remain a regular meat eater after reading the content from panel A and a 28% chance they would intend to be a low meat eater after reading the content from panel A. Conversely, for those with the same background who were exposed to control there is a 47% chance they would remain regular meat eaters and a 15% chance they would intend to be low meat eaters. In short, exposure to panel A almost doubles the probability that a regular meat eater with no intention of changing habits actually reports to want to lower their consumption by two degrees (from regular to low) compared to the control group (.283 vs .154). We see similar trends across the different panels relative to control which corroborates our findings from our linear models with binarized outcomes.

4 Conclusions, Study Limitations and Future Recommendations

Our study attempts to understand how the development of different forms of evidence would impact on intended meat consumption. Our findings from our first randomization suggests that all three (population health evidence, personalized nutritional evidence, and environmental evidence) may have a significant impact on reducing intent for meat consumption however there was limited signal that any specific type of evidence was more likely than others to drive this intended reduction in meat consumption.

Within our second randomization, all significant effects seen in the first randomization with the exception of population health evidence (Panel A) were no longer significant. The most likely explanation for this is a "demand effect" resulting from the study subjects increasing understanding of the objective of the experiment given that they had already completed a similar exercise on a different panel for their first randomization. The consequence of this is that a larger proportion of survey subjects provided a response suggesting they will eat less meat for all panels - *including the placebo panel X* - in the second randomization, thereby diluting most of the treatment effects seen in the first randomization.

There are several limitations of our study which could benefit from simple design changes if this study were to be repeated

1. As noted in Section 3.1.1, there was a different rate of compliance between those who were assigned the placebo panel versus those who were not. If this experiment were conducted again, we would recommend adjusting the placebo format to include a text-based passage rather than only an image review. This would have resulted in more similar experience to our treatment panels and may have better identified non-compliers to ensure a true

apples-to-apples comparison. However, given that the current design required a survey subject to view two panels (and to correctly respond to the comprehension check on both panels), this was not considered a major limitation in identifying compliers for our analysis.

2. Our compliance rate could have been improved if we had structured our MTurk survey incentives in such a way to ensure more thoughtful responses. This could have been achieved by providing a small base pay (e.g. 10 cents) for survey completion, but an additional payment (e.g. 30 cents) for those who have demonstrated thoughtfulness in their responses. This could have been a combination of answering the comprehension check question (task 1) correctly as well as providing some free text (task 2) which indicated review of the passage. This would have given us greater confidence that the responses to the outcome variable (task 3) was meaningful rather than just a rapid click-through exercise.
3. The randomization procedure did not occur precisely as intended, given that Qualtrics ensured a balanced randomization only over the combined two panels rather than for both the first and second randomizations separately. This was a known limitation of using Qualtrics which we considered an appropriate trade-off relative to the risk of high attrition seen in the pilot where we controlled the randomization procedure ourselves by splitting the survey into two parts. The two alternatives we would explore moving forward are (1) to only collect information on one panel for each survey subject to ensure complete randomization for each of the four panels or (2) collecting information on all four panels for each survey subject in a randomized order. However, given the demand effects observed in our second randomization, it is possible that the meaningfulness of each subsequent panel reviewed by the same study subject would suffer from similar issues, and therefore we would likely recommend focusing on option (1).

The treatments considered in this analysis can be considered complete, in that they each provide a unique hypothetical situation, distinct from each other. This means that we cannot distinguish the treatment effect associated with the specific nuances of each panel, which uses different wording and images to characterize the evidence we are interested in assessing. Whilst this is relevant to highlight, this was very much a feature of our study question, as we wanted to understand how three distinctly different evidence types would impact on propensity to change meat consumption. Future studies which may be interested in understanding the impact of specific messaging should consider alternative panel designs to capture the impact of nuanced messaging.

Finally, it is important to discuss what the impact of our findings would be on actual meat consumption. A study such as this which focuses on the hypothetical suffers from two limitations if we want to translate these findings to what will drive an actual reduction in meat consumption: (1) hypothetical judgements are hard to make, and each individual's threshold for the evidence level required to judge whether a statement is “true” is hard to know, and (2) an intention to change meat consumption may be very different from the actual realized behavior, given the challenges involved in changing individual behavior. Despite these limitations, behavioral change must start with an intent to change and this has been addressed in this study.

5 References

Celis-Morales, Carlos *et al*. Effect of personalized nutrition on health-related behaviour change: evidence from the Food4Me European randomized controlled trial. International Journal of Epidemiology, Volume 46, Issue 2, April 2017, Pages 578–588.

United Nations Intergovernmental Panel on Climate Change (UN IPCC). Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems (August 2019). https://www.ipcc.ch/report/srccl/

World Health Organization. Q&A on the carcinogenicity of the consumption of red meat and processed meat (October 2015). https://www.who.int/features/qa/cancer-red-meat/en/

6 Appendix

The following provides a full replication of our survey.

Baseline Survey prior to randomization

My gender is:

- Male
- Female
- Other

My age is between:

- Age 18-25
- Age 26-35
- Age 36-45
- Age 46-55
- Age 56 or Over

My household income in USD is between:

- $25,000 or Less
- $25,000 - $49,999
- $50,000 - $74,999
- $75,000 - $99,999
- $100,000 or More

The highest degree or level of school I have completed is:

- Some High School, No Degree
- High School Degree
- Bachelor's Degree
- Master’s Degree
- Doctorate Degree

During a typical week, I am a:

- Regular Meat Eater (2 or more meals a day with meat)
- Moderate Meat Eater (1 meal a day with meat)
- Low Meat Eater (less than 1 meal a day with meat, some days I don’t eat meat at all)
- Vegetarian (no meat, but some animal-related products)
- Vegan (no meat or animal-related products)

I am currently:

- Trying to eat more meat, milk and/or eggs
- Trying to eat less meat, milk and/or eggs
- Happy with my consumption of meat, milk and eggs

Four Randomized Panels - each participant assigned to see two sequentially

Panel A: (Population Health Example)

Please review the paragraph below and answer the following questions assuming that the content is true:

Scientific research proves that eating meat in general will increase the risk of dying from cancer, heart attacks and strokes, and accelerates the aging process. Eating high levels of red and processed meats results in worse health outcomes, and even consuming white meat and animal-related products (dairy and eggs) increases the risk of poor long-term health.

1. Based on the above passage, eating meat will (select the best answer):
    a. Decrease your metabolism
    b. Increase your appetite
    c. Increase your risk of heart attacks and strokes
2. Please provide a few comments below on how you would feel about the above passage. [free text]
3. If the above passage were true, then during a typical week I would prefer to be a:
    a. Regular Meat Eater (2 or more meals a day with meat)
    b. Moderate Meat Eater (1 meal a day with meat)
    c. Low Meat Eater (less than 1 meal a day with meat, some days I don’t eat meat at all)
    d. Vegetarian (no meat, but some animal-related products)
    e. Vegan (no meat or animal-related products)

Panel B: (Individual Health Health Example)

Please review the paragraph below and answer the following questions assuming that the content is true:

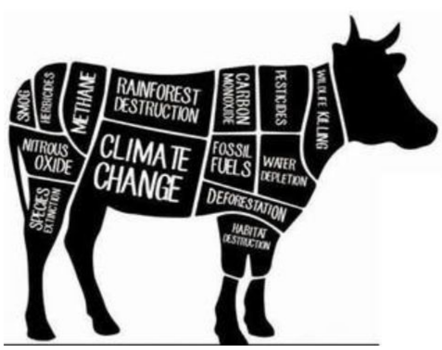
Imagine that in the near future, your annual check-up now includes a genetic screening via DNA sequencing as well as a physical examination, behavior assessment, family history, and blood tests. The combination of this genetic technology and big data can provide an FDA-validated personalized assessment of your health risks associated with your behavioral profile. Your results come back that eating meat (and to a lesser extent meat-related products like dairy and eggs) will increase your risk of premature death from cancer, heart attack and strokes and will accelerate your aging process.

1. Based on the above passage, eating meat will (select the best answer):
    a. Increase your risk of premature death
    b. Benefit your overall health
    c. Help build a health database
2. Please provide a few comments below on how you would feel about the above passage. [free text]
3. If the above passage were true, then during a typical week I would prefer to be a:
    a. Regular Meat Eater (2 or more meals a day with meat)
    b. Moderate Meat Eater (1 meal a day with meat)
    c. Low Meat Eater (less than 1 meal a day with meat, some days I don't eat meat at all)
    d. Vegetarian (no meat, but some animal-related products)
    e. Vegan (no meat or animal-related products)

Panel C: (Environmental Health Example)

Please review the paragraph below and answer the following questions assuming that the content is true:

The global scientific community has proven that climate change exists and that livestock is the primary driver due to their massive emissions of climate-changing greenhouse gases. They find consumption of beef in America alone creates the equivalent greenhouse gas emissions of 50 million cars. While red meat causes the greatest strain to the environment, white meat also causes significant impact as well. This research makes it evident that if more people shift to a vegetarian or, better yet, a vegan diet, we could make a meaningful impact on the direction of climate change.



1. Based on the above passage, livestock is responsible for (select the best answer):
    a. improving gas mileage
    b. increasing greenhouse gases
    c. increasing flora and fauna
2. Please provide a few comments below on how you would feel about the above passage. [free text]
3. If the above passage were true, then during a typical week I would prefer to be a:
    a. Regular Meat Eater (2 or more meals a day with meat)
    b. Moderate Meat Eater (1 meal a day with meat)
    c. Low Meat Eater (less than 1 meal a day with meat, some days I don’t eat meat at all)
    d. Vegetarian (no meat, but some animal-related products)
    e. Vegan (no meat or animal-related products)

Panel X: (Placebo)

Please review the images below and answer the following questions:

1. What do you see in the above images?
    a. Cars and Trains
    b. A sky filled with stars
    c. Baby Animals
2. Please provide a few comments below on how this makes you feel. [free text]
3. If the above passage were true, then during a typical week I would prefer to be a:
    a. Regular Meat Eater (2 or more meals a day with meat)
    b. Moderate Meat Eater (1 meal a day with meat)
    c. Low Meat Eater (less than 1 meal a day with meat, some days I don't eat meat at all)
    d. Vegetarian (no meat, but some animal-related products)
    e. Vegan (no meat or animal-related products)