Measuring Change in Diet for Animal Advocacy

Jacob Peacock

Humane League Labs Report E009R01

Published November 7, 2018; Revised November 12, 2018

Abstract

We consider the self-reported dietary measurement instruments currently recommended for animal advocacy research and possible alternatives to using these instruments. Reviewing the biases associated with self-reporting and comparing these self-reports with direct biochemical measurements of diet, we conclude the animal advocacy research community should strive for more accurate measurements of diet change and move away from self-reported measurements. Having identified this problem, we provide an overview of current and developing alternative measurements of diet change. First, we review commercial sources of food purchasing data. Second, we consider efforts to collect data with retailers and institutions through relationship-building with those organizations. Third, we consider predictive biomarkers as a developing alternative for directly measuring the diets of individuals. Finally, we relate the needs of animal advocacy researchers to the broader issues of measuring and understanding the food system.

This work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/ or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.
1 Introduction

Many interventions in animal advocacy seek to reduce the demand for animal products in order to reduce the number of animals used for food. In particular, interventions like leafletting, online advertising and pledge programs often urge individuals to adopt more plant-based diets. Intervention research seeking to measure these changes has thus far almost exclusively used self-reported outcomes, whereby individuals are asked to complete a survey on their dietary behaviors or intentions. In this report, we examine the biases associated with self-reports of change in diet and the data validating these self-reports against biochemical measures of diet. In light of this evidence, we explore alternatives to self-reporting in hopes that popularizing these alternatives may lead to more animal advocacy research using non-self-reported measures of diet.

Before considering the details of self-reported diet change and its alternatives, we should consider more precisely measurement goals of intervention research. Ultimately, intervention research seeks to measure the change in the number of animals required for use in food caused by an intervention. Of course, the number of animals affected is not readily measured and instead several proxies may be used. Self-reports of diet are designed and validated to measure the amount of food actually eaten by an individual. By contrast, the commercial and collaborative data sources explored in this report measure food disappearance, which includes the significant fraction of food that is never consumed due to waste, spoilage or other loss. The prevalence of dietary patterns, like veganism and vegetarianism, is another proxy to consider. However, interventions encouraging individuals to eat fewer animal products (termed “reducetarian”, “flexitarian” or “semi-vegetarian” diets) are widespread in animal advocacy and could represent significant change in diet. As such, measuring the incidence of dietary patterns may be insufficient to evaluate these interventions and this proxy is not further considered here. While no measurement is perfect, intervention research should strive to use proxies that can be used to approximate the effect of interventions on animal lives.

2 Self-reported dietary measurements

Both Animal Charity Evaluators [1] and Faunalytics [2] have promulgated best practices for self-reported dietary measures using food frequency questionnaires (FFQs) and the Automated Self-Administered 24-Hour (ASA24) dietary assessment tool. In an FFQ, the respondent is asked to indicate the frequency with which they consumed particular food items over a period of time or their intention to consume those food items in the future. In the ASA24, respondents are asked to comprehensively recall the last 24 hours of their diet through a
Both approaches are prone to significant biases, both intrinsic to survey research and unique to their use in animal advocacy research. Examining validation data comparing actual and self-reported diet confirms these biases are strong in self-reported dietary measures.

### 2.1 Biases

All surveys are subject to response bias and those used for diet measurement are no exception. FFQs rely either on respondent recall, which is prone to memory error, or respondent’s prediction of their future behavior, itself unreliable. ASA24 mitigates the risk of memory error by using a shorter recall period; however, this is at the expense of higher variance since the diet of the last 24 hours might not be representative of the respondent’s average diet. In either, approach pre- and post-treatment surveys can be employed to compare the respondent’s diet before and after an intervention. Additional measurements can further reduce variance and provide a better estimate of long-term consumption behaviors. However, this approach introduces non-response bias if some respondents do not follow-up and the possibility of reactivity, where the repeated measurements cause respondents to alter their behavior.

Social desirability bias, where respondents are inclined to provide answers they believe to be socially desirable or desirable to the researcher, is a particular concern to animal advocacy research. Animal advocacy interventions often rely on arguments surrounding the negative moral, health, environmental and social justice implications of consuming animal products. As such, it is reasonable to believe social desirability bias in self-reports is large. While propensity for social desirability bias can be measured and then mitigated via statistical modeling, without validation one cannot be assured of the success of this approach.

### 2.2 Validation

With these biases in mind, researchers have worked to validate self-reported diet measures against recovery biomarkers, a small group of molecules that can be used to biochemically measure some aspects of diet. In particular, sodium, potassium, protein and energy consumption can all be measured with high accuracy. By comparing the self-reported and direct biochemical measures of diet, we can understand the errors introduced by self-reporting.

Freedman et al. collated the results of five validation studies measuring protein and energy consumption. Each study used a combination of food frequency questionnaires and multiple 24-hour recall surveys. While the ASA24 is not directly tested in any of the vali-
dation studies, its progenitor, the Automated Multiple Pass Method (AMPM) developed by the United States Department of Agriculture, is used in four of the five. The AMPM is administered by a trained interviewer with five rounds of questions and careful prompting on serving size; it represents an ideal self-reporting of diet. As such, we might expect the accuracy of the AMPM to be an upper bound on that of the ASA24.

The results of this analysis are reported in terms of the Pearson correlation coefficient, which represent the strength of correlation between two variables, with zero indicating no correlation and one perfect correlation. Overall, these coefficients are low. For energy consumption measured using the combined results of three 24-hour recalls, coefficients averaged across the studies were 0.28 for men and 0.34 for women. FFQs fared worse, with coefficients of 0.16 and 0.24 for men and women, respectively. For protein consumption, the numbers are somewhat better with coefficients of 0.48 and 0.49 for men and women using 24-hour recalls and 0.28 and 0.29 using FFQs. However, even in the best cases, the self-reported dietary measurements explain less than a quarter of the variance in actual diet.

These results also imply an increase in variance resulting from the use of self-reports (See Section). This increase in variance reduces the statistical power of intervention research using self-reported outcomes, necessitating a larger sample size. To illustrate this, consider an experiment evaluated with a two-sample t-test, aiming to detect a medium effect size of 0.5 with 80% power and false positive rate $\alpha = 0.05$. Self-reporting as above with 24-hour recalls would increase the number of required participants 1.4-fold in the best case scenario of measuring protein consumption in women. In the most extreme case, 5.2 times as many participants are required to measure energy for men using self-reports rather than direct measurement. Especially in mass outreach like leafletting or online advertising, where a small effect may still be of interest, this loss of power poses a barrier to successfully detecting effects. Given these results, which use more rigorous survey instruments, applied in triplicate, and without significant social desirability bias, we should assume the self-reported surveys common in the animal advocacy literature are even less accurate.

3 Commercial marketing data

A more direct alternative to self-reporting, is measuring what consumers purchase. In the United States, there are three major streams of consumer food purchasing data available to researchers: retail scanner data, measuring the sales across retailers; the National Consumer Panel, surveying a panel of consumers to measure their purchasing; and customer loyalty...
cards, which track the purchases of individual consumers at particular retailers. Comprehensive retail scanner data and access to the National Consumer Panel is available through two companies, Nielsen and IRI, while smaller companies handle scanner data for just a few retailers (eg, 84.51) or specialized sectors (SPINS for organic and natural foods). [10] Both data sets are available for purchase through negotiation with either company or, for academics, at a significant discount through a partnership with the Chicago Booth School’s Kilts Center for Marketing. [11, 12] The market for customer loyalty card data is still developing, with numerous vendors representing different retailers. These sources and their strengths and weaknesses are described below, with weaknesses shared by all three explored in a final section.

3.1 Retail scanner data

Retail scanner data is collected at the point of sale (ie, cash register) of 35,000 stores across the United States, describing each item purchased by consumers with a Universal Product Code (UPC) and the purchase price. The data available from Nielsen or IRI is aggregated across regions and product types, although this is likely negotiable. However, the academic data set is available disaggregated, but anonymized at the retailer level, and starting in 2004. While scanner data includes a wide variety of food retailers, some major retailers do not participate (eg, Trader Joes and Aldi). [12, 13] Comparison to other data on US retailers and their sales suggests IRI retail scanner data covers only 41% of retailers and 55% of food and alcohol sales dollars, although some of this discrepancy is attributable to different inclusion criteria of retailers. [14] Furthermore, private label products, which are provided and sold by only a particular retailer, are often omitted from retail scanner data or aggregated in such a way as to impede the calculation of prices per unit. [14]

3.2 National Consumer Panel

The National Consumer Panel (NCP; formerly known as the HomeScan Consumer Panel) is a joint project between IRI and Nielsen whereby a panel of households is recruited and incentivized to use a bar code scanner to report all items purchased each week. [15] About 40-60k individuals are represented in the panel, with data beginning in 2004 and tracking the same set of products as the retail scanner data. [11] While this method ultimately relies on participants voluntarily scanning items, recall and prediction biases associated with self-reported measures are mitigated and social desirability bias would at least require more direct action (ie, deliberately not scanning items).
However, the NCP is less suitable for intervention research. First, the panel is already under intense scrutiny, with surveys and “special studies” routinely run on the panel. These may exacerbate reactivity and render results non-generalizable. Second, the panel has a retention rate around 50%, suggesting a strong selection bias, and may not succeed in being representative of the US population. Third, products sold by weight (so-called “random weight products”) may not have the associated weights recorded in the data set, producing another uncertain step in analysis. Fourth, some animal advocacy interventions are not readily applied to individuals spread across the country; for example, billboards or even direct mail would be difficult to distribute to panel participants cost effectively. That said, alternatives to the NCP are under development, especially those reducing the burden on participants by using optical character recognition (OCR) to rapidly scan receipts, rather than arduously scanning item by item.

3.3 Customer loyalty card data

Customer loyalty cards are issued by retailers to their customers and often include discounts and other rewards for purchasing products at that retailer. Loyalty cards can then produce longitudinal data on the detailed purchasing habits of a particular customer at that retailer. However, these programs are still subject to selection and reporting bias, where only some customers opt to enroll in a loyalty card program and may not always use their card. That said, anecdotal reports suggests even data from a single retailer is well correlated to national aggregate trends.

3.4 Common challenges

Despite the “Universal” of UPCs, coding and accuracy is still a concern as many UPCs are not readily matched to their corresponding products. Other UPC codes need to be anonymized so as to maintain the anonymity of retailers carrying private label products. Loyalty card data produced via OCR may not even include UPCs.

Unlike self-reported measures of diet change, commercial marketing data does not include all food (consumed) away from home (FAFH), which accounted for 34% of calories in US diets in 2013-14. (However, some FAFH, like a meal at a friend’s home, may still be included in grocery scanner data.) Additionally, scanner data is available only for brick and mortar stores, not online retailers, although Nielsen and IRI have separate data products covering some of these channels.

Further complicating matters for animal advocates, it is crucial that changes in diet

DOI 10.17605/OSF.IO/8ZQC3
be translated to number of animals effected, which necessitates estimating the fraction of animal products in some items. This data is not readily available and is at odds with the internal organization of the marketing-oriented data vendors, which cater to providing data on individual products, rather than all products containing chicken meat, for example.

4 Direct collaborations with restaurants, grocery stores and institutional dining services

Given the challenges and expenses present in commercial marketing data, exploring more direct collaborations with restaurants, grocery stores and dining services at colleges and universities is a developing avenue for obtaining data on dietary change. While the exact data available, feasibility of forming such relationships and costs are less clear, at least one group of researchers has had important success. By collaborating with a dining hall at Occidental College, Joshua Tasoff and collaborators procured “a rich dataset of individual-level food consumption data of three meals per day, 20+ weeks per year for the ~2,000 students at the college.” [21]

Such data would be more powerful than that provided by retail scanner data, which cannot track changes across individuals. College campuses are a particularly valuable source of data as they tend to be geographically isolated, with students consuming many of their meals and spending much of their time on campus. This allows interventions to be readily applied and measured. Furthermore, the college demographic has always been a prime target for animal advocacy intervention as college students tend to have agency in their food choice and be eager to explore new world views. The Menus of Change University Research Collaborative is actively working to facilitate collaborations between researchers and universities in the US, with sustainability goals that align relatively well with animal advocacy outcomes. [22]

Such collaborations could also be forged with public institutions like schools and hospitals or anywhere that serves food to individuals. Academics and researchers at non-profits may be especially well positioned to build mutually beneficial good-will collaborations. In forthcoming research, Humane League Labs will be working to expand collaborative programs to other US university and college campuses, which could provide data of unprecedented quality to animal advocacy researchers.
5 Predictive biomarkers

As discussed in Section 2.2, recovery biomarkers are available to precisely measure a handful of dietary parameters. Recovery biomarkers are immediate byproducts of the parameters they measure and are thus less prone to variation between individuals. Predictive biomarkers are a developing approach that may offer measurements of other aspects of diet like animal product consumption. Predictive Biomarkers for animal products in the diet are usually substances produced during the metabolic breakdown of food, called metabolites, and found in blood, saliva or urine. Food additives and contaminates, like antibiotics in cows and polychlorinated biphenyls (PCBs) in fish, also have potential as biomarkers. While dietary biomarkers have been studied for decades, the field of metabolomics has recently accelerated development by measuring the levels of thousands of different metabolites and considering the patterns in which they occur. Using metabolomics, numerous predictive biomarkers correlating with different foods, food groups and dietary habits have been discovered.

Distinguishing diets like veganism, vegetarianism and meat-eating has been relatively successful, while efforts to measure consumption of dairy and eggs and meat products from chickens, cows and pigs are still in early stages. Red meat has received the most interest for its role in disease, and we might expect red meat biomarkers to be among the first developed. Meanwhile, biomarkers for animal products from many different species, like fish products, will likely be developed later. The development of dietary biomarkers must also overcome the numerous confounding genetic and lifestyle factors, like metabolic rate, smoking, exercise, and differences in cooking conditions. Current approaches also tend to focus on short-term biomarkers indicative of consumption only within the few days before testing, rather than consumption over weeks or months. With these complications and others, estimates of food consumption from biomarkers may be too imprecise to be useful for some time. However, with active coordination by the Food Biomarkers Alliance, this field is rapidly developing and perhaps in the future animal advocacy researchers could utilize such biomarkers in measuring diet change.

6 Conclusions

We’ve explored current and developing alternatives to self-reporting of dietary data. Obtaining this crucial outcome data still poses a barrier to animal advocacy research. The challenge is not unique to the field and a National Academy of Sciences report also found “gaps that limit the information that policy makers and researchers have to address current and emerg-
ing issues in food and nutrition.” [p4][18] These challenges, however, may not be intrinsic, but appear to be the product of lobbying by the food and marketing industries. For example, the Argus Leader, a South Dakota newspaper, has been embroiled in a 7-year lawsuit to obtain data on the use of the publicly funded SNAP program, which offers nutrition assistance to millions of Americans at the annual cost of $68 billion taxpayer dollars. In the last months, lobbyist from the Food Marketing Institute have successfully thwarted the release of data describing SNAP purchases from 2005-2010, arguing the decade-old information would cause competitive harm. [31]

In light of these efforts, animal advocates should align themselves in the fight for greater transparency in the food system, not just in how animals are used for food but also in how marketers work to perpetuate demand for animal products. In addition to these efforts, researchers should work to build collaborations with independent and conscientious business partners to establish direct relationships for collecting dietary data. Lastly, we should keep an eye to technological developments that may aid our research and advocate for the development of tools well-suited to animal advocacy research.

7 Technical appendix

In Freedman et al. [6] Web Appendix 3], true and self-reported diet for a particular study (k) are related using a random effects model

\[ T_i = \lambda_0 + \lambda_1 Q_i + u_i \]

where,

- \( T_i \) is the unobserved true dietary consumption of individual \( i \)
- \( \lambda_0 \) is the intercept for all individuals
- \( u_i \) is the random intercept for individual \( i \), assumed independent and normally distributed with mean zero
- \( \lambda_1 \) is the regression coefficient termed the attenuation factor
- \( Q_i \) is the observed self-reported consumption of individual \( i \)

The Pearson correlation coefficient for \( T \) and \( Q \) is then,

\[ \rho_{T,Q} = \frac{\text{Cov}[T, Q]}{\sqrt{\sigma_T^2 \sigma_Q^2}} \]
where $\sigma^2_T$ and $\sigma^2_Q$ represent the variance of $T$ and $Q$ respectively and Cov denotes the covariance. To calculate $\sigma^2_T$, using the properties of variance and the independence of $u$ and $Q$, we find:

$$\sigma^2_T = \text{Var}[T] = \text{Var}[\lambda_0 + \lambda_1 Q + u] = \lambda_1^2 \sigma^2_Q + \sigma^2_u + 2\lambda_1 \text{Cov}[u, Q] = \lambda_1^2 \sigma^2_Q + \sigma^2_u.$$

To calculate $\text{Cov}[T, Q]$, using the properties of covariance and the independence of $u$ and $Q$, we find:

$$\text{Cov}[T, Q] = \text{Cov}[\lambda_0 + \lambda_1 Q + u, Q] = \lambda_1 \text{Cov}[Q, Q] + \text{Cov}[u, Q] = \lambda_1 \sigma^2_Q.$$

Substituting and rearranging the definition of Pearson correlation coefficient, we have the result presented in Web Appendix 3:

$$\rho_{T,Q} = \lambda_1 \sqrt{\frac{\sigma^2_Q}{\lambda_1^2 \sigma^2_Q + \sigma^2_u}}.$$

As we are given $\lambda_1$ and $\rho_{T,Q}$ (Tables 6 and 7, respectively), we can recover the ratio of the variance of self-reported and true consumption,

$$\left(\frac{\rho_{T,Q}}{\lambda_1}\right)^2 = \frac{\sigma^2_Q}{\lambda_1^2 \sigma^2_Q + \sigma^2_u} = \frac{\sigma^2_Q}{\sigma^2_T},$$

and then the ratio of standard deviations,

$$\frac{\rho_{T,Q}}{\lambda_1} = \frac{\sigma_Q}{\sigma_T}.$$

To examine the effects on power, consider a $t$-test and the associated Cohen’s $d$ effect size of the form $\Delta/s$, where $\Delta$ is the difference in means between the two treatments and $s$ is the shared standard deviation in both treatments. If we could measure consumption directly, we might detect an effect size $d_T = \Delta/\sigma_T$. However, if we used self-reports instead, we’d find an effect size, $d_Q = \Delta/\sigma_Q$. (This is ignoring issues of bias, which can be readily corrected, so we assume $\Delta$ would remain constant.) By dividing $d_T$ by $\sigma_Q/\sigma_T$, we can relate $d_Q$ and $d_T$:

$$\frac{d_T}{\sigma_Q/\sigma_T} = \frac{\Delta}{\sigma_T} \frac{\sigma_T}{\sigma_Q} = \frac{\Delta}{\sigma_Q} = d_Q.$$

Thus we can calculate the reduction in effect size resulting from using self-reported rather than direct measures of diet. For the Python code to perform these calculations, see the Open Science Framework repository at [https://osf.io/8zqc3/](https://osf.io/8zqc3/).

DOI 10.17605/OSF.IO/8ZQC3
8 Revision history

- November 12, 2018: Non-substantive copyediting.
- November 9, 2018: Non-substantive copyediting.

9 References


[12] *Nielsen Datasets Retail Scanner Data*. URL: [https://research.chicagobooth.edu/nielsen/datasets#simple2](https://research.chicagobooth.edu/nielsen/datasets#simple2) (visited on 08/21/2018).


DOI 10.17605/OSF.IO/8ZQC3


DOI 10.17605/OSF.IO/8ZQC3
