Measuring Food Intake in Field Studies

Brian Wansink
Cornell University

April 21, 2008


* Brian Wansink is the John S. Dyson Chair and Professor of Marketing at Cornell University, Department of Economics and Management, Address: Cornell University, Department of Economics and Management, 110 Warren Hall, Ithaca, NY 14853-7801, USA. Phone 607 255 5024, Email: wansink@cornell.edu
Measuring Food Intake in Field Studies

I. Introduction – “Cool Data”
   A. The Emerging Importance of Field Studies
   B. A Framework for Defining, Measuring, and Analyzing Intake

II. Studying Intake
   A. Reducing Systematic Variance through Intake Process Models
   B. Increasing Effect Sizes by Identifying Boundary Conditions
      1. Intake-Prone Foods.
      2. Intake-Prone Populations
      3. Intake Floors and Ceilings

III. Selecting Intake Measures
   A. Pre-Intake Measures
      1. Acquisition Measures.
      2. Measuring Intake Intentions.
   B. Post-Intake Measures
      1. Measures of Intake Recall.
      2. Residual & Inferential Measures of Intake.
   C. Scenario-based Methods of Measuring Intake

IV. Analyzing Intake Studies
   A. Screening and Segmenting Nonusers
   B. Measuring Usage Frequency
      1. Identify Broad Intake Intervals.
      2. Specifying Intake Situations and Usage Occasions.
      3. Accounting for Usage Facilitators.
   E. Measuring Intake Volume (per occasion)
      1. Actual vs. Inferred Intake.
      2. Counting vs. Weighing.
      3. Accounting for Satiation.
      4. Identifying the Appropriate Level of Analysis

V. Conclusion
Measuring Food Intake in Field Studies

Abstract

An increasing number of studies are investigating food intake volume and frequency in field studies outside the lab. Yet in this emerging research context, inconclusive and seemingly contradictory findings may be the result of complications in defining, measuring, and analyzing intake. This chapter offers a framework for conducting intake studies, and it outlines methods and measures that will increasingly lead to successful studies and accurate conclusions.
I. Introduction – “Cool Data”

When I was a Ph.D. student, I heard two words that changed my research life.

At the time, a brilliant but under-published assistant professor was given tenure at Stanford. He had published only three articles at the time. While this would be unheard of at most other research schools, the grapevine indicated that his tenure success was partially due to the overwhelming impact of these three articles. Shortly after his tenure announcement, I saw this professor at a campus running track. After a few laps of small talk, I asked him about his secret to research success. For the last 20 years, his two-word answer has influenced me more than anything I learned in my Ph.D. program.

His secret to research success? “Cool data.”

Cool data. That summarizes most of what many of us do not want to do. We tend to be experts at lab studies, complex modeling exercises, or short-term trials involving begrudging sophomores who need the extra credit. The last reason we joined academics was to run around in restaurants, movie theatres, bars, and dining rooms. Yet this is where some of this “cool data” hides. It is data from real people in real situations who are being observed, coded, measured, and dispassionately analyzed and reported (Schachter 1971).

Cool data is hard to collect. It can be data collected in restaurants (Bell and Pliner 2003), sports bars (Wansink and Cheney 2005), household pantries (Terry and Beck 1985), college cafeterias (Levitsky and Youn 2004), food courts (Chandon and Wansink 2008), elementary school lunchrooms (Kahn and Wansink 2004), supermarkets (Iynegar and Lepper 2001; Wansink, Kent and Hoch 1998), movie theatres (Wansink and Park 2001). It can also be data collected from unusual populations, such as preschoolers (Fisher, Rolls, and Birch
2003) or amnesiacs (Rozen et al, 1998), or which contrasts data collected on the streets of Paris (Wansink, Payne, and Chandon 2007) with that collected on the streets of Philadelpia (Rozin et al, 2003).

It is harder to get IRB approval for it, it is harder to set up, it is harder to staff, and it is harder to analyze (Meiselman 1992). Yet what I have also learned is that cool data can capture imaginations, it can suddenly make science relevant to an unsuspecting segment, and it can almost always be published, eventually.

Over the years, I have struggled with many types of cool data. Some has been unpredictable, some has been costly, and some has ended up being influential. I conducting over 200 studies, I have estimate this:

- 40% of the studies are disasters and need to be rethought because of a conceptual mistake, or rerun because of a methodological mistake.
- 40% of the studies come out differently than hypothesized, leading to a new understanding, a new theory, and a new confirmatory study.
- 20% of the studies come out perfectly predicted as hypothesized.

My favorite insights, and therefore my favorite studies, have almost always come from the second category. Part of this success has been because I have strived to set up studies in a way that generate diagnostic results. But that is only part of the key. Part of what distinguishes great papers from no paper is the ability to analyze field data to find the needle in the haystack. The purpose of this chapter is to provide some of these insights in how to set up cool data studies and how to analyze them to extract something that is true, insightful, and generalizable.
While the intent of using controlled field studies is to make our results powerful, impactful, and interesting, this does not always happen. What follows are some of the methods to try and mistakes to avoid.

A. The Emerging Importance of Field Studies

Until recently, intake research developed by borrowing constructs from parent disciplines (primarily economics and psychology) and modifying them to suit questions related to intake behaviors (see Bruner 2002 for a review). Recently, however, an emerging issue that is more uniquely central to intake – that of food intake volume and frequency – has become of increasing importance to consumers and public policy officials (Allison et al 1999; French, Story, and Jeffery 2001). Perhaps because there is a less established set of intake measurement procedures, methods, and analyses, the contributions in this area have been more sporadic and less consistent than in other areas. The objective of this chapter is to provide a framework that helps improve probability of success and the consistency and synergy of intake-related field studies.

At this stage of development, there are many unexplained inconsistencies in intake studies. Some show that package shapes do not cause increases in intake (Raghubir and Krishna 1999), while other studies do (Wansink 1996). Some show intake to be driven by cognition (Berry Beatty, and Klesges 1985), while others show it driven by perceptions (Inman 2001). Some show that attitude is related to intake, while others show it is not. We contend that many of these inconsistencies can be attributed to a wide variance in the methods, data analysis approaches, and reporting practices of intake studies. This chapter
offers a framework for conducting intake studies, and it outlines methods and measures that would most likely to lead to successful studies and accurate conclusions.

There is a basic distinction between “choosing” and “using” that makes research difficult to conduct without a framework. One way intake is distinct from choice is because it involves a behavior that can sometimes be more consequential than choice while also being less deliberate (Wansink and Sobal 2001). Because intake has immediate implications for a person’s gratification or even health, issues of intake frequency or volume are not always explained by rational economic models (Assunsao and Meyer 1993). Importantly, there is also a motor component that distinguishes usage and intake studies from pencil and paper studies or even from those involving the selection and purchase of foods off of shelves (Schacther and Freidman 1974).

B. A Framework for Defining, Measuring, and Analyzing Intake

The differences noted above can complicate the ability to investigate intake using standard models, methods, measures, or analyses. Such differences generate uncertainty as to whether weak effects in intake studies are due to noise introduced by the measure of the method, or whether they are instead due to the effect itself is genuinely insignificant. A key objective here will be to increase effect size while decreasing systematic variation. There is a wide range of contexts in which this variation can occur. They can occur in studies of how the size, shape, variety, structure, or inventory level of a food influences intake. It can also occur in studies of how different advertising strategies influence intake frequency either in new situations or as substitutions for existing foods. They may also include situation variables
such as how do social norms, the presence of others, or occasion-based usage influence intake differently (Clendenne, Berman, and Polivy 1994; Fuenekes, de Graf, and van Staveren 1995).

There will be a number of benefits to a framework that improves the consistency in the way researchers use models, methods, measures, and conduct analyses. First, such a framework will help researchers determine the most fruitful and sensitive way to study an intake-related variable without having the results obscured by unnecessary noise or inappropriate analysis. Second, it will help the area of consumer research grow in a more structured, synergistic manner by helping provide better structure to methods, measures, and analyses. Last, it will importantly encourage process models to be developed that can be useful in further understanding intake by helping examine it in an un-confounded, least invasive manner.

As shown in Figure 1, we will first examine how intake can best be studied by increasing effect sizes and decreasing noise. Next the timing of pre- and post-intake measures will be discussed. Last, analysis techniques that are somewhat unique to intake studies will last be investigated. This paper examines intake primarily in the context of the usage volume and frequency of consumables.

II. Studying Intake
Intake can be examined by using lab experiments, field studies, surveys, or consumer panels or it can be examined less directly by projecting intake from purchase data or by inferring intake from what is discarded. The specific research question dictates what method is most appropriate and can provide the strongest effects with the least interference.

The basic objective in developing a framework of intake measurement sensitivity is to increase relative effect sizes by reducing the systematic variation. One important way to accomplish this is to eliminate or control the potential sources of unwanted variation or noise associated with a study. Reducing systematic variance can be accomplished by developing intake models that help one articulate these differences. Developing a process model for intake experiments allows you to specify details and variables which your hypothesis involves (Baron and Kenny 1986; Evans and Lepore 1997). It allows for the examination of variables which could potentially obscure results or influence intake and also considering alternate explanations, you can eliminate or add variables to to decrease ambiguity, and factors which could obscure hypothesized results.

For instance, consider the seemingly ubiquitous American phenomenon of cleaning one’s plate. Research involving this tendency can simply show that it happens under different conditions and at different ages (Rolls, Engle, and Birch 2000. These would be effect studies. Process studies would also try to explain why this happens (Wansink, Payne, and Werle 2008). For instance, it might occur because they amount served to a person represents an implicit “consumption norm” (Wansink 2004) of how much is appropriate, typical, normal, and reasonable to consume. Therefore, one almost dutifully complies. By specifying this mediating variable (“the amount on my plate is the typical amount a person should eat”), one
can begin to explain why an effect occurs. Effect studies explain “what,” and process studies explain “why.”

As will be explained in the next section, hypothesizing the different reasons why you think an independent variable might influence a dependent one, can add a lot of precision to your field studies and to their analysis. One big reason for going to this extra effort is that it can help open up the “black box” of why people do what they do. A second reason it is can help trouble shoot results that do not seem consistent with expectations or which seem too noisy to draw a conclusion.

In one study of Pacific combat veterans we investigated how one’s involvement in World War II influenced how much they would eat when they were introduced to a Chinese buffet dinner 55 years later (Wansink 2006, pp. 154-156). The initial results were disappointing and showed no difference in comparison to a control group of the same age.

What we had not done is to fully articulate the reason why we thought they would eat less Chinese food. Our unarticulated reason was that we thought the negative feelings they had about their wartime experience in the Pacific would bias their long-term preference toward Asian food. Our implicit process model is outlined in Figure 2.

As noted, results of this study were disappointing. There was no difference between the veterans and a control group. Having a process model enabled us to trouble-shoot the data and determine whether the theory was wrong or whether there was another explanation. Of
those veterans who enjoyed Chinese and Japanese food and still ate it with some frequency, there were no characteristics they had in common. Before the war, some had lived in big cities, some on farms. Some had graduated from college, others had never seen a 9th grade classroom.

What did explain their behavior was the level of combat they had experienced as soldiers. When analyzing the profiles of those Pacific veterans who liked Chinese food, we did not find Marines who had been at Iwo Jima or infantry soldiers at Guadalcanal. What we found were mechanics, clerks, engineers, and truck drivers – enlisted men who did not experience the War from the front line. Although their wartime experience was a sacrifice, they did not come home with terrible associations that tainted the taste of food even 50-60 years later.

Having a process model enabled an analysis of the data that yielded insights other than what were expected. Even though our process model was wrong, it gave us a framework to analyze the data more completely. The subsequent insights were more on track with what is really true about the formation of long-term eating behaviors. This is one reason that developing a process model is an important way to start the planning of a field experiment.

A. Reducing Systematic Variance through Intake Process Models

Increasing the relative effect size can be accomplished by using the appropriate method and by either controlling or eliminating those noise-related variations that may obscure hypothesized results. Many factors can influence intake and obscure the effects of a study. One way to determine those that would be most damaging to a study is to specifically
articulate the process model that drives one’s resulting hypotheses. This is important even for effect studies or calibration studies. Even though the focus of such studies may not be on the process leading to behavior (intake), per say, articulating such a model can help researchers to eliminate sources of unwanted variance (such as the time of day, the presence of others, the cost or availability of an item, and so on) or to control them.

For instance, suppose one is studying how lighting level influences how much a person eats. In order to estimate the most accurate influence of lighting, it would be important to statistically eliminate the effects of other intervening variables, such as social facilitation (what others at the same table are doing). By measuring and accounting for the intake of other people at the table, a researcher can have a more pure examination of the influence of lighting (de Castro 1994).

Regardless of whether one is studying the influence of variety, serving sizes, atmospherics, or structure on intake, it is important to hypothesize why intake should occur. Developing process models suggest not only the level at which to analyze data, and what potentially confounding variables should be measured, but it can also suggest which foods to use in the study (Rolls, Bell, Waugh 2000). Doing so will help control other variables or processes that might be seen as either confounds or as potentially explanatory causes. Even when these extraneous factors are not controllable, insights from a model can suggest measures that could serve as covariates or would enable post-hoc tests to be conducted.

Explicitly articulating an intake model can help organize a study. Furthermore, doing so suggests important factors that must be controlled for and key measures which must be taken in order to best understand the process that is occurring. As an example, consider the process model illustrated in Figure 3. In this model, the relationship which suggests a process
model where two mediating factors – consumption norms and consumption monitoring – are suggested to explain why five common environmental cues have such an empirically robust influence on intake (Wansink 2004).

The process of doing this enables a researcher to articulate why hypothesized effects will occur, but it will also help them either measure, control, or eliminate those factors that could mask or eliminate the hypothesized relationship. The impact of package size on intake, for instance, can be explained by scarcity research, by perceived differences in cost, or by a perceptual explanation (Wansink 1996). Articulating these alternatives can help a researcher design studies which either rule out or account for alternative explanations. It is important to show when a effect does work as well as when it does not. This helps other researchers to conduct more efficient follow-up studies and it makes their efforts less derivative (Lawless et al 2003). As a result, researchers can investigate the fundamentals behind intake without having to conduct derivative or incremental research based on replications and slight extensions of existing studies.

Generating a process model also helps specify “boundary” conditions by knowing when and how effects will be found (Pliner 1974). In this way, different antecedents will yield different results. One example of how an understanding of how boundary conditions influenced the analysis of many studies involves eating restraint.
Understanding that there was a difference between restrained and unrestrained eaters disclosed many effects that were previously hidden (Polivy et al 1979; Polivy et al 1986).

**B. Increasing Effect Sizes by Identifying Boundary Conditions**

By specifying a process model that relates how various factors influence intake, we can develop a better understanding of when an effect may or may not be observed. For example, in their study of pouring and intake behavior, Lee and Raghubir (2002) found that less involved consumers were more influenced by the shape of packaging compared to more involved consumers. Given these conclusions, it is important that future studies related to package shape make certain their subjects are highly involved and are not distracted during the study. Not only do boundary conditions help develop stronger manipulations, they also show conditions where the standard set of manipulations can be more effective than others.

1. **Intake-Prone Foods.** Some foods are more prone to intake acceleration or to variation in intake volume than others. Neslin and Ailawadi (1999), for instance, found that intake acceleration occurs with certain foods (such as yogurt) but not with others (such as ketchup). In this case, the intake of the more convenient and hedonic foods (yogurt) was more sensitive to stockpiling pressures than was ketchup (which less conveniently requires being eaten with a prepared food). Making these food-level distinctions is important because researchers can sometimes over-rely on existing databases when testing intake-related hypotheses. If the existing categories on which there is data (e.g., coffee, ketchup, yogurt and so on) are not categories that are prone to intake variation, however, researchers would be unfairly handicapping their hypotheses.

Recent studies have found that foods can be most acceptable to intake variations when they are hedonic and convenient to prepare and consume (Chandon and Wansink 2002).
Furthermore, it might also be that widely but irregularly consumed foods would be promising to examine as intake-prone categories. In contrast, foods that are widely enjoyed (such as coffee) but are consumed in reasonably regular patterns may not show the variation in intake that is necessary to provide an accurate examination of hypotheses.

2. Intake-Prone Populations. In attitude, perception, and cognition research, there is an understandable bias toward using the data from all participants who have been involved with a study. The rational is that everyone has attitudes, perceptions, and cognitive processes that must be accounted for to accurately assess the influence of the different experimental treatments.

Measurement sensitivity in intake research, on the other hand, presupposes that users are different and potentially of more interest than non-users. Since the key questions are focused on intake volume or frequency, it may not make sense to include participants who never have used (or who will never use) the food. For instance, a study using milk as a food stimuli can justly screen out participants who are lactose-intolerant or otherwise allergic to milk. Similarly, it would also be acceptable to screen out non-users of milk or people who simply do not drink it because they do not like its taste or texture. Studies, which do not screen out non-users are left with columns of zeros because regardless of the treatment condition, the consumer will not consume the food. While this may not be a problem if these non-users are balanced across all cells of a design, this still generates unwanted variance that can be justly eliminated.

Yet just as nonusers add unwanted variance because they are qualitatively different than users, there may also be occasions when heavy users may be qualitatively different than light users. For instance, those with an alcohol dependency might behave qualitatively differently in a wine study than an occasional wine-with-dinner person, and Americans who have a “sweet
tooth” behave qualitatively differently in a dessert study than most Asians, who have a stronger preference for less sweet, more savory snacks.

It is important to understand that screening heavy users is only justified when they are believed to qualitatively differ than lighter users. In general, it is still important to account for past usage when examining intake or usage behaviors. Past usage behavior can be used as a covariate for various analyses, or it can be used as a basis for categorizing heavier users from lighter users, and analyzing each group separately.

3. Intake Floors and Ceilings. One difficulty with emerging fields of research lies not only in defining and measuring the dependent variables, but also in determining the appropriate range of the independent variables that are examined. That is, in order to determine how different package sizes might influence one’s intake volume, we must know the general range of sensitivity toward package size. If the small size package is such that most people empty the package on each intake occasion, it is a confounded test of the hypothesis because participants could not consume more even if they had wanted. It is therefore important to select packages that are large enough to not be fully consumed.

While the naive solution would be to increase the relative size of both packages to the point where there would be no ceiling effects, doing so would also result in diminishing effects. In his study of package sizes, Wansink (1996) showed that there becomes a point at which larger packages cease to have an effect on how much people consume. In effect, there is a point at which a person is just too full to eat (or drink) any more, regardless of how much stronger the manipulation becomes. If the manipulations are set outside this range of intake sensitivity, the resulting conclusions will be that the stimulus has no impact on intake. One solution to this involves calibration prestudies. Another involves collecting data at three or more reasonably
wide intervals of strength, knowing that only two may end up being of eventual interest because of either floor effects or ceiling (diminishing return) effects.

A related concern when examining quantities is the problem with satiation or flavor/category burnout. Inman (2000) identified that the burn-out that is related to satiation can be delayed when a person switches between flavors within the same category. This can be one way in which to delay the effects of burnout in order to more extensively examine intake. Yet this underscores, more generally, the importance of understanding that satiation can cause a natural ceiling that we cannot easily observe. One solution for this is to be mindful of the time horizon under which a study is being conducted. The longer the time horizon, the less of a concern for burnout.

III. Selecting Intake Measures

Intake-related variables can be measured on various levels. This includes acquisition (such as purchase) and intended intake (which can be either short-range or longer-range), continues to actual intake, and can conclude with post-intake reporting or residual analysis. Realizing that there are various levels on which intake can be examined becomes important in determining the most effective (and cost-efficient) way to study a research question. Figure 4 illustrates the different points at which intake can be measured. These can generally be characterized as pre-intake measures or post-intake measures. If a research question does not necessitate the most involved measure (about intake, for example), it may be easier and less expensive to collect measures of intake intention instead.
It is important to distinguish between these different stages. It is often incorrectly assumed that one stage necessarily and accurately leads to another (Weingarten 1984). For instance, it is often assumed that purchase will predictably lead to intake, and that purchase data, therefore, is a surrogate for intake. While perhaps true under some conditions, there are still many foods that are thrown away or which sit unused and forgotten in the back of cupboards and pantries.

A. Pre-Intake Measures

1. Acquisition Measures. As noted earlier, the intake rate of foods between purchase periods may be linear or consistent. While this may be true for some foods like coffee or toothpaste, most other foods are less consistently consumed. Consider over-the-counter medicine, such as cold medicine or cough syrup. Such medicines are not consumed at a linear rate. Some consumers buy and consume this medicine immediately, while other consumers deplete it in tiered stages whenever they are sick. Still others stockpile the food for “insurance,” while others purchase it and never use it because they either forgot about it or misplaced it. Scanner data does not tell us who in the household consumed the food, and it often assumes a household size of one.

While the drawbacks of these have been noted earlier, there do seem to be conditions where they may be more justifiable. With frequently purchased goods, there is less likelihood that extraneous events (such as parties, or guests) are causing untractable spikes in intake. With nonperishable items there is less of a likelihood that things will be thrown away without
being consumed. Last, when items are widely consumed around the household, it may be that they will be consumed at a more static rate than if there is only one person consuming the food.

2. Measuring Intake Intentions. The assumption of much marketing and sensory research is that if consumers rate a flavor as acceptable, they will consume it. Yet basic measures of attitude do not explain behavior, and even less frequently relate to behavior (Garber et al 2007). Because purchase and intake is a marketing-related objective in food development, intake intentions should be measured at the time of testing. Two ways of measuring one’s intake intentions (for a particular time period, such as “within the next two weeks”) are through likelihood measures, or through estimates of one’s future intake frequency. Likelihood measures can be directly obtained by asking an individual how likely (“Highly Unlikely” = 1 to “Highly Likely” = 9) it will be that he or she consumes the food within an upcoming time period. Intake intentions can also be measured by asking one to estimate how many times he or she might possibly consume the food within a similar time period (Wansink and Ray 1992).

These two different measures of intake intent have different relative strengths. With infrequent users of a food, frequency estimates will be skewed toward 0 units (especially over a relatively short period of time). This is partially a drawback of numerical estimates that provide no gradation between 0 and 1 unit. In such cases, the frequency estimates provide less variance
and less information than an estimate of intake likelihood. With light users, intake likelihood estimates will provide greater gradation in response and more sensitivity in detecting any potentially different effects a particular set of sensory qualities would have on intake.

In contrast, with frequent or heavy users of a food, a frequency estimate is likely to be more accurate than a likelihood estimate. This is because the distribution of these frequency estimates is more likely to be normally distributed. As a result, a frequency estimate of one’s intake intent is likely to provide more variance and more information about the intended intake of heavy users than is a likelihood measure, which would undoubtedly be at or near 1.0 (100 percent probable). With heavy users, frequency estimates would be a more accurate estimate of a heavy user’s future intake frequency of a food.

When a sample consists of both heavy and light users, both likelihood and frequency measures should be used with both groups. However, in weighting the relative measures, frequency estimates should be weighted more heavily for heavy users consumers, and purchase likelihood measures should be weighted more heavily for light consumers. Using both measures allows some degree of comparison, but weighting them allows more confidence in making segment-level conclusions.

In general, however, intake intention measures are most valid when they involve a readily accessible food that involves little or no preparation. The more intermediate steps that are involved (such as purchasing the food or preparing the food), the less accurate this measure becomes. In general, the direction of the bias for intake intention measures (frequency and likelihood) depends on the availability and convenience of the food.

When conditioning on past usage it is important to check for differences across groups and foods. For frequently consumed foods, it may be that the difference between relatively
heavy and relatively light consumers may not be important or diagnostic. Conversely, with very infrequently consumed foods, there may not be much of a difference between relatively heavy and light consumers because these differences might actually be caused by differences in consume volume per occasion and not by a wider number of occasions each.

The relation of intentions to usage is conditioned on many factors. It is critical to ask about intake in a specific situation. This is consistent with the work of Fishbein and Ajzen (1980) who argued that attitudes are always specific to a situation and the key to linking attitude to behavior was to make sure that the attitude measure was carefully and specifically linked to behavior in a actual situation. The importance of linking attitude to specific situations is illustrated by the corresponding correlation increase; estimates for general intake in three food categories had a correlation near .45; estimating the intake in specific situations increased the correlation to .77.

B. Post-Intake Measures

1. Measures of Intake Recall. It has long been recognized in social psychology that environmental factors can influence intake. One such confound involves those around us. There are social norms as to how much should or should not be consumed. These can involve social norms of under-intake at an interview dinner, or social norms of over-intake at a New Years party.

When studying unfamiliar foods, for instance, it is generally assumed that people have similar food experience backgrounds and that the treatments will have a general impact across all consumers. In contrast, with intake-related studies, one’s past intake of the food and their liking of the food can easily create biases that need to be accounted for so they do not
generate unwanted noise. For instance, with studies of how stockpiling influences intake, the
treatment conditions may involve giving some participants high inventory levels of specific
foods while giving others lower levels. In such circumstances, it is important to account for
existing inventory levels and to account for any additional purchases that might occur during
the test period.

On an immediate level, one’s like or dislike of a food influences their intake. This is
something that can be controlled by using past measures of attitude toward the food or toward
the general category of foods it represents. Having people with a wide range of preferences
for tuna fish or for chocolate can influence the results. Some of this would presumably be
taken in to account when screening or segmenting non-, light-, and heavy-users. In other
cases, people do not have the opportunity to eat foods with the frequency they like (because of
availability or diet restrictions). Yet their inherent preference for them would suggest they are
different from light or non-users who exhibit similar intake tendencies.

Estimated recall of intake can also be made and analyzed. While this is a reasonable
surrogate for actual intake, it is influenced by factors such as the salience and the convenience
of the food. Recent studies of chocolate intake have shown that people consume more
chocolates when it is convenient and visible (on the desk) than when it is either inconvenient
(2 meters from the desk) or convenient but not visible (in the desk). Yet their estimates of
how much they eat were in the opposite direction. If the food was convenient, they
overestimated how much they consumed, whereas if it was inconveniently located, they
underestimated their intake of it (Painter, Wansink, and Hieggelke 2002).

2. Residual & Inferential Measures of Intake. Just as pre-intake predictions about
intake can be made forward from a purchase, post-intake inferences can be made based on
residual waste, including what a person does not eat or what a person throw’s away (“garbology”).

Intake can be measured by how much people use, how much they do not use, or how much they remember using. Perhaps the easiest way to measure intake is by weighing or counting the what is not eaten; this entails pre-measuring portions and then weighing the food that was left uneaten.

One study, which measured the impact that container size and mood had on the intake of movie popcorn, gave people either large or extra-large containers of popcorn (Wansink and Kim 2005). Prior to passing these out, each was weighed and the weight was written on the bottom of each container. Following the study, each container was collected and their weight of the remaining popcorn was subtracted from the original weight to provide an estimate of intake. One danger of measuring residual intake is that it must be noted if people either spilled part of the container, whether they combined it with a friend, or whether they emptied the container and filled another container (or their pockets) with the food for later use. In all of these cases, the data from these people is eliminated from the study.

In general, when measuring usage, it is much more effective to weigh rather than count. This is true when giving foods to people and also when assessing the residual. Counting out exactly 300 jelly beans is much more time consuming than making sure every container has the same number of grams of jelly beans. Following the study, determining the residual is much more easily done with a scale than with counting.

C. Scenario-based Methods of Measuring Intake

One emerging method is that of scenario-based studies. Such studies present consumers with a scenario and ask them to predict what their intake would be under various manipulated
circumstances. Scenario-based studies can be an effective means of investigating intake issues, particularly in the preliminary stages of a study or when beginning a program of research. While this is not as conclusive as having actual measures of intake (or of reported intake), it can be effective as a pre-study. This has even been used in early stages to determine how the size of containers would influence how much people poured. In these studies consumers were given the stimuli (various sized bottles of shampoo), and asked to indicate how much they would use by drawing a circle on one photograph of a hand. The area of the circle was then used as a surrogate for volume and used in the subsequent measures.

Intake studies conducted using scenario-based or hypothesized intake can be appropriate for academic research. Other times, however, it is also important to try and understand how such scenarios relate or map to actual intake so the size of the effect can be calibrated to determine whether it is worth intervention. This can be accomplished by recreating the study in a field situation. As the number of these field replications grows, a more general estimate of the carry-over of estimation and hypothesized intake measures can be made.

IV. Analyzing Intake Studies

In analyzing intake studies, there are critically different methods used to analyze intake that differ depending on whether it is volume, frequency, or incidence that is being studied. While each hold their own challenges, we will focus first on the practice of screening or segmenting non-users and then address each in turn.
A. Screening and Segmenting Non-users

Non-users can create unwanted variance in intake-related studies through non-reporting or through reactivity. Screening non-users out of an analysis can be critical in reducing variance that could otherwise obscure the effects through non-reporting. This is frequently been done when studying volume, and it can also be justified when studying frequency.

Yet when studying incidence, screening out nonusers can be problematic. It would be easy to misclassify a person as a “non-user” if too narrow of a time window or too restricted of a range of situations were investigated. There are many reasons a person might choose to not consume a food. It may be occasion-based (such as only eating cranberry sauce during holidays), or it might be joint-decision-based (such as being indifferent to anchovies but willing to eat them if ordered by a spouse or colleague). In either of these two cases, it is important to realize that there are differences between non-incidence that is caused by an absence of intake opportunity versus non-incidence that is caused because one strongly dislikes a food and never consumes it.

Determining whether a participant should be screened out as a non-user can present problems. Screening a person’s food intake prior to a study can unintentionally foreshadow the purpose of the study and contaminate the results. One way to avoid this is to post-screen respondents instead of pre-screening. This involves asking the screening questions (past usage frequency, most recent intake, and so on) following the collection of the measures for the primary dependent variables. At this point, reports of their past usage will not bias the primary measures, and consumers who simply do not use the food can be safely selected out of subsequent analyses. While nonuse of this food would result in data that would be thrown away, the screening can be done without fear of any contamination or demand effects.
There are times, however, when prescreening must be done for cost or efficiency reasons. In these cases, one less reactive way in which to prescreen consumers for a study is to ask them their past usage frequency of the food while also asking them about their frequency of a number of other related or unrelated foods that act as distracter items. Screening questions could be asked which do not directly investigate the target food but collect enough related data for an effective prescreen.

Separating consumers into different groups (such as non-, light-, and heavy-users) can be useful in determining the influence of marketing variables. Conditioning on past usage is particularly important when examining intake intentions and estimated intake.

Just as screening out non-users can increase measurement sensitivity by reducing noise, sometimes focusing only on heavy users can increase sensitivity by magnifying the potential effects of certain treatments. One way to prescreen on this basis is by the absolute past frequency in which they have consumed a food.

In a study of how advertising variations influence intake, it was found that heavy-users were more influenced by intake-oriented ads than light-users, but were more sensitive to treatment methods. As a result, differences that would be statistically insignificant for lighter-users would be significant to heavy users.

B. Measuring Usage Frequency

1. Identify Broad Intake Intervals. It is important to not examine intake during potentially confounding time periods when the base rate of intake of the food is too low to be diagnostic (soup tests in the summer in Texas; cranberry sauce tests during the spring). It is critical that either diary panels or callbacks match the intake cycle interval. While some foods
(such as coffee or toothpaste) are consumed at fairly standard rates, many others (particularly seasonal non-perishables) are not. Intake should be sampled across a long enough time frame to not be biased by accelerated intake that might occur during the early part of the purchase cycle, or to be biased by the decelerated intake that might occur at later stages. The former bias would lead to a long-term overestimate of intake, and the latter would lead to a long-term underestimate.

It is also important to differentiate between intake driven by special occasions versus intake which occurs in a more standard way during the typical passage of time. Chandon and Wansink show that it is important to differentiate between intake that is determined by salience vs. that driven by convenience.

2. Specifying Intake Situations and Usage Occasions. Although purchase intentions are often used as a dependent variable, intake intentions have a less developed history. Because of the differences between intake and purchase, there are legitimate concerns that intake intentions are not reliable surrogates for actual intake. Many of these concerns are valid when intake intention measures are incorrectly taken, however, when carefully and correctly done, their reliability and validity improves.

As with attitude measures, intake measures should specify the time period and the intake situation. For instance, Wansink and Ray (1996) report that when people were asked to estimate their intake of a food in a specific situation (eating dinner within the next two weeks), they were more accurate than when estimating their total intake of soup over the same time period \((r = .77 \text{ vs. } r = .58)\).

One way to validate this relationship between intentions and intake is to conduct a split-half study. Such a study measures estimated and actual intake with one group (before
and after being exposed to various stimuli) and measures only the actual intake of a second group (after being exposed to the same sets of stimuli). Comparing the two groups enables a researcher to assess potential demand effects. It also allows the researcher to determine how reliable intake intention estimates are in that situation by comparing intended and actual intake. Should they prove to be accurate and if they are not reactive (no demand effects), future studies of this type might be more efficiently done by simply using intake intentions.

3. Accounting for Usage Facilitators. Recall that certain foods might have different intake acceleration propensities than others because of physical characteristics such as whether they are hedonic or utilitarian. It is also important to note that there are wide ranges of situations that can influence intake and must also be accounted for. Recent studies indicate (Chandon and Wansink 2002) that convenient-to-eat foods (such as ready-to drink iced tea or pre-made pudding) have a much higher intake incidence than do foods that are less convenient to consume (such as ice-tea mix or boxed pudding). Similarly, the basic convenience and salience of a food also influences how frequently it is eaten. A recent study showed that chocolate kisses that are more accessible and convenient (sitting on the desk) are eaten with greater frequency than those that are slightly less accessible and convenient (two meters from the desk) (Painter, Wansink, and Hieggelke 2002).

The importance here is to try and account for environmental factors that can influence intake. This can be accomplished in either the design of a study or by controlling them by asking critical questions about convenience, salience, or accessibility. It seems that people feel a need to take a food’s visibility and convenience into account when they try to estimate their prior intake of it in such a way that a food that is inconvenient to consume may be eaten in larger amounts than are thought or recalled. Similarly, dietary researchers need to take into
account the visibility and convenience of foods because not doing so might lead to biases in intake recall studies and diary panel estimates. One way to allow for such biases is to ask research participants to rate the visibility and convenience of the foods under investigation; these ratings can then be used either as covariates or as blocking or segmentation variables.

4. Comparative vs. Absolute Measures of Usage Frequency. Intake frequency is typically measured on an absolute level – units per week. It can also be measured, however, on a more comparative basis using either difference measures or an index. This would be akin to accounting for individual-level variation by accounting for with-in subject effects. Such difference measures can involve the difference in intake between the test period and a pre-test calibration period. They can also involve the differences in intake between the test period and one’s usual or typical intake of the food over that same time interval. This can be expressed as a difference measure, but they have also been expressed as an index.

Another means of assessing intake is through the use of scaled comparison measures. Questions which address this typically ask a person to compare their intake with that of a typical intake period. They might then rate this on a nine-point scale (1 = I consumed less than is typical for me; 9 = I consumed more than is typical for me). While such comparative measures lack the appeal of an actual, tangible unit of measure, they appear highly sensitive to the treatments. They may be useful in pre-studies as well as in final studies as a validation check or self-report in follow-up studies.

C. Measuring Intake Volume (per occasion)
Measuring intake volume is more straightforward than measuring intake frequency, yet still important. The issue of screening out non-users is critical here as with intake frequency.

1. Actual vs. Inferred Intake. The two general methods for calculating intake are measuring actual intake volume or measuring residual intake which is inferring intake based on the difference between the amount taken initially and the amount left after consumption.

A variation on these methods measures how much food is taken and assumes everything is consumed unless there is residual. A recent study of pouring behavior examined how glass sizes and shapes influenced pouring volume of juice at a summer camp. When people returned their dishes, the residual juice was subtracted from the amount originally poured. The difference was assumed to be what was consumed.

While this is a generally accepted measure, it needs to be underscored that it is inaccurate if a person spills or shares a food, or if they save it for a later time. To account for these, the person can be given an exit questionnaire that asks whether they spilled, shared, or saved the food in a pocket, purse, briefcase, or backpack. These people can then be eliminated from the study and reported as such. If it is critical that the study remain as unobtrusive as possible, unobtrusive video cameras can record intake behavior and spilling, sharing, or saving can be coded and associated with the relevant person. This entails individual intake or assigned seating.

2. Counting vs. Weighing. Two different methods of determined intake volume are to either count or to weigh the residual amount of the food that is left at the end of the intake episode. This can either be subtracted from what was originally taken by the consumer or by
what was given to the consumer. In some instances, foods can be weighed, but in other instances it is preferably from a design or a reporting standpoint to count the units consumed.

Weighing what is consumed is nearly always preferable to counting what is consumed. Indeed, with liquids and more continuous (mashed potatoes) or non-uniform consumables (such as French fries or meat), weighing can be the only option. Unfortunately, keeping track of what foods are taken and consumed based on weighing can be obtrusive and impractical for some studies. In these cases it can be more advisable to discretely measure how much is taken either through unobtrusive observation, or by recall questionnaires.

Certain foods lend themselves better to being counted (vs. weighed) than others. With more discrete or uniformly shaped foods, it may be preferable to measure or at least report them in terms of units. This is often true with candy (such as small candies or small candy bars), snacks (crackers, grapes, or cheese squares), or other small pre-prepared foods. This has also been done successfully by asking the number cans of a soft drink, or glasses of a beverage one has consumed during a time period. The importance of counting foods is that it can often be done less obtrusively than weighing. That is, people can be observed by confederates, they can be recorded by unobtrusive video cameras, or they can be questioned at the end of the episode.

Yet in other cases, it may be expedient to weigh a food in grams, but report in it terms of units (or percentage of units). This can be done by weighing the food in the study, and translating this measure to units. For instance, in studies of M&M and jelly bean intake, Kahn and Wansink (2004) used a residual weight measure in the methodologies, but translated these weights in to the number of M&Ms and number of jelly beans consumed for ease of reporting.
3. Accounting for Satiation. Among others Rolls (2002) and Inman (1999) showed that a key issue in examining intake was factoring in the effect that satiation from additional foods might have on the intake of the primary food being studied.

For example, in one study of candy intake (Kahn and Wansink 2004), consumers were given different varieties of candies that only varied in their color. The objective was to determine how variations in color influenced intake volume. Although carefully done across multiple locations with multiple groups, no differences were found. The methodology entailed bringing consumers into a central mall location, giving them their choice of beverage and assigning them to one of the candy conditions where they were given candy to eat as they watched television. It reviewing the methodology, it was later found that their selection of a beverage influenced how much candy they ate. Consumers who selected a sugared soft drink, ate less candy than those who selected a diet soft drink, and both groups ate less than those selecting only water. When the study was rerun using only water, the expected differences in intake were consistent with what was hypothesized (Inman 2001).

To make certain the effect sizes are as large as possible, it is important to determine how the presence of other foods influences intake. A recent field study we investigated how the presence of others (manipulated through table size) influenced salty snack intake was conducted in a pub environment. We did not expect people to drink additional beverages than the ones we provided and carefully controlled. As a result, a good portion of the study was confounded because it did not also take in to account the amount of beer that people drank. That is, while larger tables increased the amount that people eat, it also increased the amount they drank, and it is well-established that the two variables are corollated.
It is also important in these studies of volume to take measures of critical covariates that can influence intake. Besides basic gender, age, height, and weight variables, it is also good to ask how many hours since their last meal, and a rating scale as to the extent they are on a diet and watch what they eat. These measures should be used as covariates, and in the early stages of exploration, they can be assessed or analyzed separately to determine the strength of an effect. Not surprisingly, unless a ceiling is reached, the strongest differences between treatments are found among males in their early twenties who do not carefully monitor what or how much they eat.

4. Identifying the Appropriate Level of Analysis. One issue concerning intake is to not lose sight of the forest because of the trees. It is important that researchers consider a wide range of intake when analyzing how different factors influence intake. Consider a recent study that examined whether people consumed more olive oil or butter when eating bread dinner at a restaurant. While it was found that people ate many more calories of olive oil than butter on a piece of bread, they ate fewer pieces of bread. As a result, people do use more olive oil than butter when they eat a piece of bread, but they also eat fewer pieces of bread. By broadening the field of focus, examining the amount of bread consumed, the results of the study became more conclusive.

When people are given olive oil in a restaurant, they consumed 26% more on each piece of bread than those given block butter but they ended up consuming 23% less bread in total. This finding provides a novel look at a traditional dietary recommendation: When considering the health aspects in a diet, one needs to focus not only on fat calories, but on total calories. By eating less fat, one can unknowingly eat more carbohydrates -- eating twice as many low-fat cookies trades off fat calories for an increased amount of carbohydrates.
When focusing on the intake of a target food, it is critically important to also analyze how that food influences the intake of related foods. A preoccupation with fat intake can distract health care professionals from understanding alternative effects which may bring their own set of health concerns.

Similarly, recent studies investigated how the degree of heterogeneity in candy assortment influenced the intake of candy. In conducting the study, people were given one of the various assortments of candy the amount they ate was observed while they watched television. Because candy can make one thirsty, they were also given a selection of beverages. The data were inconclusive, and it was suspected that those who chose bottled soft drinks (instead of bottled water) might have varied how much candy they ate in order to keep their total sugar intake at a reasonable level. A subsequent study that provided only bottled water to people supported this conclusion and the results were more in line with what was originally predicted.

V. Conclusion

“Cool data” is not easy data to collect. When it comes to field studies the challenges of measuring intake every procedure and set of measures has both its advantages and disadvantages. As the need for precision and accuracy increases, increased confidence in intake methods and measures can be found by combining methods and triangulating on their conclusions. For instance, combining a lab study with a field study can provide both the tight internal validity in the lab study with the external validity of the field study. Similarly, it can
be useful to combine secondary data with recall surveys, or to combine diary panels with laboratory experiments.
References


Levitsky, D., & Youn, T. (2004). The more food young adults are served, the more they overeat. *Journal of Nutrition* 134:2546-2549.


Table 1.
Different Forms of Intake-Related measures

<table>
<thead>
<tr>
<th>Usage Level</th>
<th>Measures of Intake (within next two weeks)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Likelihood Measure</td>
<td>Volume Measure</td>
</tr>
<tr>
<td><strong>Light Users</strong></td>
<td></td>
<td>.421</td>
<td>.151</td>
</tr>
<tr>
<td>-- Canned Soup</td>
<td></td>
<td>.779</td>
<td>.580</td>
</tr>
<tr>
<td>-- Gelatin</td>
<td></td>
<td>.490</td>
<td>.221</td>
</tr>
<tr>
<td>-- Cranberry Sauce</td>
<td></td>
<td>.472</td>
<td>.161</td>
</tr>
<tr>
<td>-- Average</td>
<td></td>
<td>.190</td>
<td>.618</td>
</tr>
<tr>
<td><strong>Heavy Users</strong></td>
<td></td>
<td>.233</td>
<td>.462</td>
</tr>
<tr>
<td>-- Canned Soup</td>
<td></td>
<td>.043</td>
<td>.207</td>
</tr>
<tr>
<td>-- Gelatin</td>
<td></td>
<td>.197</td>
<td>.597</td>
</tr>
</tbody>
</table>
Figure 1.

Different Forms of Intake-Related Variables

Defining Intake

- Reducing Systematic Variance Through Process Models

- Identifying Boundary Conditions
  -- Intake-prone products
  -- Intake-prone people
  -- Avoiding floors and ""

Measuring Intake

- Pre-intake Measures
  -- Acquisition
  -- Intake intention

- Post-intake Measures
  -- Intake recall
  -- Inference

- Scenario-based intake Measures

Analyzing Intake

- Screening and segmentating nonusers

- Analyzing Intake Frequency

- Analyzing Intake Volume

- Analyzing Intake Incidence
Figure 2.

How World War II was Believed to Influence the Eating Habits of Pacific Veterans

Figure 2A. The Original *Incorrect* Process Model

The World War II Experiences of a Pacific Veteran → Long-term (55 years later) Negative Feelings Toward Memories Associated with the War → Negative Feelings Toward Chinese Food and Reduced Intake at a Buffet (compared to a control group)

Figure 2B. The Revised Process Model

The Level of Combat Experienced by World War II Pacific Veterans

The World War II Experiences of a Pacific Veteran → Long-term (55 years later) Negative Feelings Toward Memories Associated with the War → Negative Feelings Toward Chinese Food and Reduced Intake at a Buffet (compared to a control group)

High Level of Combat → Long-term (55 years later) Neutral or Positive Feelings Toward Memories Associated with the War → Favorable Feelings Toward Chinese Food and Increased Intake at a Buffet (compared to a control group)

Low Level of Combat
A Process Model of How the Food Environment Influences Consumption or Intake Volume

**The Food Environment (The 5 S’s)**
- Salience of Food
- Structure and Variety of Food Assortments
- Size of Packages and Portions
- Shape of Serving Containers
- Stockpiling of Food

**Consumption Norms (What We Believe is Normal)**

**Consumption or Intake Volume**

**Inaccurate Monitoring**
Figure 4.
Different Forms of Intake-Related measures

- Acquisition Measures
- Intake Intention Measures
- Scenario-based Intake Measures
- Actual Intake
- Intake Recall Measures
- Inferred Measures of Intake