2017

Perceptions of the Word “Healthy” on Moral Balancing and Licensing Patterns

Sophia Tate
Skidmore College

Perceptions of the Word "Healthy"

On Moral Balancing and Licensing Patterns

By

Sophia Tate

A Thesis Submitted to

Department of Economics

Skidmore College

In Partial Fulfillment of the Requirement for the B.A Degree

Thesis Advisor: Qi Ge

May 2, 2017

Abstract

This paper examines the complex relationship between the subjective FDA-regulated word "healthy" and its effect on consumers moral balancing and licensing habits. I create a survey to test the ability of the word "healthy" to balance or license future decisions. The results indicate that if the word "healthy" is on a package, regardless of nutritional quality, consumers use "healthy" to license unhealthy activities the next time period they are faced with a decision in the moral sphere of health. However, the balancing effect is lost after two time periods. The findings suggest that balancing and licensing patterns may depend on habit and tastes and preferences.

Acknowledgement

Many thanks to my thesis advisor Professor Qi Ge, Jiebei Luo for her help with cleaning data, my survey advisor Professor Sandra Goff, and peer reviewer Lukas Raynaud for all their help with my senior thesis. I also thank the Skidmore Opportunity Fund for their grant that helped fund my survey.

1 Introduction

In order to feel fulfilled, people try to hold themselves in high moral self-regard (Khan and Dhar 2006). However, throughout the day, people must face decisions that challenge their short- and long-run self-images; while the decisions are usually not drastic (i.e. robbing a bank or donating a kidney), a series of seemingly trivial decisions (like whether or not to tip the barista or what type of jam to buy) cause can decision fatigue (Tierney 2011). Additionally, when these decisions can challenge one's morals, they can be even more burdensome. For example, the decision between drinking coffee or tea may be an easy, guilt-free choice, as they may be acceptable substitutes for some consumers. However, choosing between eating a free donut or a free banana from the break room at work could be a more difficult decision because the donut is a treat but the banana is not. Similarly, Khan and Dhar (2006) find that after imagining completing community service, the choice between buying designer jeans (a hedonistic good) or a vacuum cleaner (a utilitarian good), calls morality into play. Imagining doing community service frames a positive self-image, and the participants in the study were more likely to choose to buy designer jeans. However, when consumers make decisions that do not align with the high moral image they subconsciously try to maintain, they also—subconsciously or sometimes consciously—try to offset their bad decision or action with a "good" one (Khan and Dhar 2006). This is known as moral balancing.

Most choices are made with other factors in mind; there are few choices truly independent from the influence of past decisions or the expectation of making future decisions (Khan and Dhar 2007). Based on past actions and decisions that build a positive self-concept, an individual may license a bad behavior; for example, if she engaged in

community service in the morning, she may feel like she can treat herself to an indulgent purchase in the afternoon (Khan and Dhar 2006, Wertenbroch and Dhar 2000). Inversely, a past guilty decision that builds a negative self-concept may require an individual to balance it with a good action or decision—if one does not pick up after her dog while on a walk, she may feel more inclined to donate to a charity.

According to classical economic thought, consumers are rational, utility maximizing agents (Mankiw 2018). That is, they have perfect, complete information that allows them to make decisions that will best maximize their utility. This theory allows economists to analyze possible scenarios of consumer behavior. However, in reality, consumers do not act in this way; they have incomplete information, and they may not always maximize their utility. For example, though the Food and Drug Administration (FDA) sets dietary guidelines for Americans to follow a healthy diet, over one-fourth of the United States population is obese (President's Council on Fitness, Sports & Nutrition 2017). Assuming that health is part of a consumer's utility function, failure to adhere to a healthy diet or get adequate exercise is not the behavior of rational consumers. Behavioral economics seeks to better understand consumer choices. In this paper, I analyze how consumers engage in moral balancing and licensing; this research is at the intersection of economics, business ethics, psychology, and in my study, public health.

Many consumers are unfamiliar with the U.S. Department of Agriculture's dietary guidelines (Office of Disease Prevention and Health Promotion 2015). Because they are unsure what is healthy or unhealthy, consumers rely on package health claims or nutrition facts. On some packages, health claims are pared down to make it easier for people to understand ("50% more protein"); the statistic may lead consumers to believe the

scientific medical community supports the statistic, and therefore they may trust in the product more. For example, a product could contain 5g of protein (10% of the daily value (DV)). If it increased fiber by 50%, to 7.5g of protein, the company could attach the new health claim to the label. While a 50% increase sounds like it may be significant, an increase of 5% of the daily value of protein has a small effect on the consumer's daily intake. This effect would be lost on consumers who rely on claims that appear to be based in science. In my study, I test to see if the word "healthy" alone was enough to balance an unhealthy decision.[1] If participants believe "healthy" was science backed, it may increase their belief in the product's ability to balance an unhealthy decision. There is a growing body of literature on consumer health behavior and willingness to pay. While the FDA ensures that no labels are false or misleading, the degree to which a label is misleading is subjective. Therefore, it is important to study consumer perceptions and behaviors so labels can provide the clear information that does not take advantage of consumers.

I designed a survey to determine how people balance their good and bad decisions in the moral sphere of health and wellness. I presented the participants with five time slots that make up an entire day and thirteen foods and activities to fill these slots with. The activities and foods were either healthy or unhealthy. Like Khan and Dhar (2007), I chose the activities and foods arbitrarily, and assigned them healthy, unhealthy, or neutral based on common ideas of healthy and unhealthy behaviors and meals. I also collected data on age, gender, location in the US, and income level. I coded unhealthy choices

[1] Packages bearing health claims are often accompanied by scientific claims on the back labels; because my survey asks participants to imagine a product with equal nutritional value but different health promotions, participants may not understand this unrealistic scenario and choose activities and foods in my survey based on habit.

as -1, healthy choices as 1, and neutral choices as 0 to act as dummy variables in my regression analysis. In addition to examining if individual participants participated in moral balancing or licensing by running regressions on their decisions after one and two lags, I also evaluate the data using analysis of variance tests (ANOVA) to determine if there is a zero net benefit from starting the day with a healthy, unhealthy, or neutral choice.

My results for the first lagged choice indicate that consumers do engage in a balancing and licensing pattern; however, with the second lagged choice, the participants were more resistant to balancing and licensing patterns. While there was not an alternating pattern of balancing and licensing (healthy, unhealthy, healthy…), participants were aware of their choices throughout the day. The sum of participants' points (based on unhealthy $= -1$, neutral $= 0$, and healthy $= 1$) was calculated, and a majority of participants had daily scores close to zero. These results of my study are valuable to the existing literature. The relevant literature in this field either examines the effects of altruistic framing on moral balancing or studies the effect of advertising health claims on consumer purchasing habits. My study is the first to focus on the effects of advertised health claims on moral balancing and licensing patterns.

I discuss more about moral licensing and balancing in the literature review and form hypotheses (Section 2). Next, I describe the methodology used to create my survey (Section 3), and then I analyze the results (Section 4). In the discussion section (Section 5), I critique my study, and in the concluding section (Section 6), I discuss policy implications and areas for further research in this field.

2 Literature Review

The phenomenon of moral balancing and licensing is well recorded in psychology—and sometimes in economics—especially pertaining to the altruistic act of donating, or being primed with a positive self-concept. In this section, I will discuss moral balancing behaviors from an economic perspective. Next, I will explain how making a health-based decision can be challenging to balance, given the complex moral space of food choices. Finally, I will discuss current advertising and FDA regulations that may be misleading to consumers.

2.1 Moral Balancing

Khan and Dhar (2006) is the seminal piece of literature on moral balancing and licensing in business and economic journals. Comprised of five studies, it highlights the effects altruism on future purchasing decisions; in each experiment, the authors frame the altruistic act differently. In one study particularly relevant to my survey, participants were divided into two conditions to examine if they would buy necessary or luxury goods. Both groups were told they had $500, but the licensing group was told the extra money was from a tax rebate. Both groups were told they could donate $100 to a charity, then they participated in an unrelated activity, and then they chose between buying sunglasses—one pair was a higher priced "hedonic" good, while the other pair was a less expensive necessity. A majority of participants in the licensing group chose the expensive luxury sunglasses.

This study effectively demonstrates the licensing effect—Khan and Dhar cite previous literature (Dahl *et al.* 2003) to determine that most consumer guilt comes from

purchasing an unessential luxury item. Pretests of consumer preferences were completed to determine how participants might value the sunglasses as either luxury or necessity. However, choosing types of sunglasses may be a biased purchase, as some consumers who wear glasses or contacts may make eyewear decisions based on their eyesight. According to a Gallup poll, over 50% of Americans primarily wear glasses, and so the decision to buy luxury or inexpensive sunglasses may have been inconsequential to them, as they would have to wear prescription glasses anyway (Newport 2000). Additional issues come from the participant pool. Undergraduate participants received credit for class—the Institutional Review Board (IRB) discourages this, citing possible conflicts of interest.

In my study, I did not frame participants' decisions with an anecdote about an act of virtue or vice to prime their next decision as in Khan and Dhar (2006) and Khan and Dhar (2007). Instead, the way participants structured their days naturally framed their decisions. For example, if participants made a healthy choice in the first time slot, that influenced their subsequent choice. In the survey I created, there were both "healthy" and "unhealthy" versions of the cereal, cheeseburger, and ice cream food choices. I presented five empty time slots and thirteen different possible choices. I gave participants two options for each type of choice: one healthy choice that participants may have viewed as necessary (e.g. cereal labeled "healthy", running on the treadmill), and one unhealthy choice (e.g. a cheeseburger labeled "delicious", watching TV), that participants may have viewed as a luxury. I also offered two neutral activities (reading and chess) that I believe are outside of the health and wellness moral space. If participants chose to add these activities to their day, they had no effect on the overall "score" for the day. While I did

not pretest consumer perception of the foods I asked consumers to choose between, it is generally accepted that cheeseburgers and ice cream are unhealthy choices, while the bran cereal in the survey may be a healthy choice. I will discuss my survey design in depth in Sections 3.1 and 3.2 on methodology and data. If participants chose the version of the unhealthy foods marked "healthy," this would have licensed them to make an unhealthy decision in the next time slot of the survey. Or, participants could have followed a different pattern of licensing, as exhibited in the following paper.

Brañas-Garza *et al.* (2013) conduct a dictator game experiment to learn more about the patterns of licensing and balancing. The authors expand on Khan and Dhar's 2006 study by using an econometric model to further explore the trends in the previous paper. It involves undergraduate students at a university in Uruguay participating in 16 sequential dictator games regarding the altruistic act of donating.[2] It is especially important to note that participants play a sequence of games with different recipients of the donation each time. Since the recipient is different each time, the authors find clear evidence of moral balancing because compensation is independent from the recipient; the compensation amount only relates to the dictator's self-image.

The authors find that the first donation amount influences subsequent donation amounts, and that there is a balancing rather than continuation effect. Additionally, donations show stability over time, as the high and low donation amounts eventually cancel each other out. The authors note that balancing and licensing does not occur at each time period; there may be a different pattern. For example, participants may feel that three good actions license a bad action, not simply one good action resulting in one bad

[2] Dictator games challenge classical economic theory, finding that external factors may influence participant altruism. Agents are expected to be rational maximizers, but this occurs at varying degrees in dictator games (Mankiw 2018).

action. From this hypothesis, they construct their model. I will discuss this in depth in the section on methodology (Section 3).

Conclusions from Brañas-Garza *et al.* (2013) are consistent with previous studies on donation and balancing (as in Khan and Dhar 2006; Khan and Dhar 2007). However, they investigate sequential decisions to understand how consumers engage in moral balancing, while the Khan and Dhar papers use time and types of goods to frame participants' decisions. This is an important aspect of the moral balancing literature; it provides evidence for the way consumers may establish routines and habits. Non-profit organizations can also use these results to maximize their donations—by understanding how people donate, they can more effectively balance their budgets and advertise to donors.

Brañas-Garza *et al.* (2013) provides a template for the methodology of my survey. Additionally, I use their model, but regress it using OLS estimates instead of GMM estimators. I also recorded participants' choices for each event, except instead of sixteen events, I provided the participants five time slots. The sequential dictator games are particularly relevant to my study, and I expect to find similar moral licensing and balancing effects. However, if participants planned a fully healthy or fully unhealthy day, indicating a continuation effect, this may have been due to habit or current health lifestyle. While I use other papers to understand how consumers conceptualize hedonistic and necessary goods, Brañas-Garza *et al.* (2013) provides a template for my survey design. I also asked participants to make a series of decisions—for my study the decisions were not altruistic, but selfish. These decisions were healthy, neutral, or unhealthy, and I expect choices to follow a balancing and licensing pattern. To

understand how consumers make selfish decisions that may require self-control, I explore another paper by Khan and Dhar (2007).

While the previous papers have investigated moral balancing when participants are primed with an altruistic act (donating, or voluntary or compulsory community service), Khan and Dhar (2007) examine how participants make decisions without altruistic priming and investigate self-control. Furthermore, the options the participants can choose from are self-indulgent and require self-control, and they are framed with future choices in mind. This study is particularly applicable to my research, as I am interested in the relationship between moral balancing and diet maintenance throughout the day, which also involves a spectrum of self-control and long-term decision-making.

One study in this paper involves female participants choosing between yogurt and a cookie for a snack. When the women were told that in their next session with the experimenters the following week, they would be given a cookie, they chose the yogurt. When the future choice was up to them, most chose to indulge in the present and balance their unhealthy choice the next week; these results are consistent with moral balancing, and they expose consumers' present-time bias. However, in another study in the paper, participants were given the choice between a high- or lowbrow film in the present, and a future choice of a cookie or yogurt, the film choice did not have an effect on the food choice; the guilt from choosing a lowbrow film did not require balancing with a healthy food choice because they are from different moral spaces.

Khan and Dhar (2007) conclude from these five studies that when participants are offered "virtues" and "vices" (which are either high- or lowbrow films, or yogurt or a cookie) with the promise of the same choice the next week, they are more likely to

choose the vice in the present. This is partially due to present-time bias, and because people are positive about their future behavior. However, personal tastes and preferences for cookies may bias the results. Participants may have previously formed perceptions of healthy food tasting bad, or they may have been lactose intolerant, or simply not liked the taste or texture of yogurt. Additionally, just because a choice induces guilt does not mean that any subsequent positive choice will balance it; choices must be within the same moral space for the licensing and balancing effects to occur.

Their study illuminates an interesting observation about health choices: when consumers are faced with a vice and a virtue, they choose instant gratification if they know they can balance their "bad" decision later. Based on this study, I hypothesize that participants in my study will license an unhealthy choice for the time slot nearest to when they take the survey because they know they can balance it in the next time period by choosing a healthy activity or food. Additionally, because exercise, sedentary activities, and healthy and unhealthy food are all in the category of health and wellness, I predict that any sequence of these choices will have balancing and licensing effects on each other. However, these habits may be too ingrained in participants' daily lives to see an effect; in the next section, I discuss the factors that contribute to the complex relationship many people have with food.

2.2 Health, lifestyle choices, and diet change

One problem physicians struggle with today is how to present patients with effective plans for diet change. Health and lifestyle habits can be difficult to change—most people cite that they do not have enough time or are too tired to workout or prepare nutritious food (Swain 2013). Understanding what makes consumers change their health

habits can have an impact on how physicians make recommendations, but also how advertisers may change their marketing strategies to increase sales.

The force of habit is another factor that influences consumer choice, but it has not yet been explored in the current literature on moral balancing and altruistic acts of donation. As Ge and Ho (2017) note, making decisions requires active, conscious participation. In order to save mental energy on making decisions, consumers form habits; these habits and subconscious activity may have an effect on public decisions. I also encountered similar econometrics problems as Ge and Ho (2017); I discuss these more in the results section (Section 4). In my study, health habits would be apparent if participants chose to follow a different pattern of choice making than the expected moral balancing and licensing effects. Because I did not collect health data on the participants, this may have been a source of bias, as different participants have varied opinions and habits on health and lifestyle.

Another factor that seems to contribute to health decisions may be national health guidelines. While the government publishes dietary guidelines for US citizens to follow in order to maintain good health, many people do not follow these guidelines. Lindsay (2010) investigates the disconnect between consumers' pursuit of health and social well-being and their understanding of health guidelines. Delaney and McCarthy (2013) add to the growing body of literature on "food moralism," concluding that consumer emotions toward food and food consumption result from the complex combination of social, physical, and economic factors that individuals experience, and that these can change over time. They also discuss the important concept of "healthism," which is the drive to improve the human body through a healthy lifestyle. Recently, the prevalence of

"healthism" is increasing, which causes consumers to attach more moral value to their food and exercise habits.

Lindsay (2010) expands this idea further. In her assessment of Australians' poor adherence to dietary and alcohol guidelines, she finds that most people are not educated about the guidelines to begin with. Furthermore, the people who are familiar with the guidelines are unsure how to follow them; this contributes to anxiety surrounding food choices. Since food and alcohol can be both sources of security or anxiety, it is difficult for consumers to intrinsically seek positive self-change and adhere to dietary guidelines.

There is a parallel between public health choices and public dictator games. Social well-being is a strong motivator for health, and many health choices are not private. Similarly, in dictator games, when the donation amount is made public, dictators are driven to donate more evenly to maintain their social well-being (Cason and Mui 1998). Changing the game from an altruistic game to one that involves other peoples' perceptions, motivates players to continue donating high amounts. Health choices and health status are relatively public.[3] For example, the fear of embarrassment that may come from checking out at the grocery store with unhealthy items may motivate unhealthy individuals to buy healthy foods. The positive feelings from eating healthy food and the shame of checking out at the store with unhealthy food items may drive the consumer to habitually buy healthy foods. Instead of a balancing and licensing effect, there may be a continuation effect because other peoples' perceptions are involved in the individual's choice.

[3] Additionally, there is a growing market geared toward making health choices more public. FitBit, LoseIt!, and MyFitnessPal have social features—users can challenge their friends to be healthier. While this may promote healthier habits, it also could create unhealthy relationships with foods, and encourage eating disorders; more research in this area is warranted before concrete conclusions can be drawn.

In sequential dictator games, making a "good" choice builds credit for making "bad" choices, and the progression is relatively symmetric, so that every high donation that is morally affirming eventually results in an equal low donation that induces guilt (Brañas-Garza *et al.* 2013). This study suggests that only the previous time period matters (period $t-1$, where $t =$ the time of the donation, so $t-1$ investigates the decision in the second period, after the first donation) for this balancing; individuals only measure against their immediate past actions. However, the decision to change a diet involves a longer timeline. An individual may decide to make healthy choices multiple days in a row and then choose one meal as the reward (may be referred to as the "cheat[4]" meal) (Jallinoja *et al.* 2010). While only the immediately previous time period allows the "cheat," the multiple day accumulation of positive "credit" to cheat may also have an effect on the intensity of the bad behavior. Additionally, the direction of the choice would be interesting to study—if a consumer knows she has a soccer game later in the day, does this license her to choose fast-food for lunch?

In my study, participants chose between a food that advertised as delicious and one advertised as healthy, though they are told at the beginning of the survey that all meal choices are nutritionally equivalent. This is similar to asking participants to choose between instant gratification and their long-term preference for health. I expect that participants would first choose the unhealthy vice option, and then balance and license subsequent options in a pattern. I am also interested in finding if the pattern is always alternating (healthy, unhealthy, healthy…) or if consumers feel that healthier choices balance an unhealthy choice (healthy, healthy, healthy, unhealthy, healthy…).

[4] However, even calling the deviation meal a "cheat" meal reinforces feelings of guilt, and may make the dieter feel emotionally worse after eating it. Perhaps calling the meal a "reward" or other positive label would be more effective.

Further study on the severity and impact of the health choice may also be interesting: do three moderately healthy choices warrant one severely unhealthy choice? How do these ratios differ across individuals? Brañas-Garza *et al.* (2013) finds that donation values even out over time, but this may not be the case with health outcomes. This could also have policy implications for future federal nutrition and physical fitness recommendations, and federal regulations of health advertising. I discuss further study in the conclusion (Section 6.2).

2.3 Advertising Health and Consumer Choice

There is an abundance of literature on advertising and consumer taste preferences. One study finds that the presence of the word "organic" on a label influences consumer preference. In Lee *et al.* (2012), participants were given "regular" and "organic" samples of cookies, chips, and yogurt; however, both samples were actually the same organic product. The researchers found that identifying a product as organic leads consumers to believe the product tastes better and are more nutritious, and that consumer's willingness to pay increases. This effect is known as the health halo. I believe that the word "healthy" by itself on a package may have a similar health halo effect.

Roe *et al.* (1999) conduct a similar study on the health halo effect, finding that it exists in the presence of health claims, too. The authors conducted short face-to-face interviews of shoppers at a mall. They presented them with mock-up food packaging designs; the designs were similar to existing brands, but no brand name was specified. The participants analyzed the designs and answered questions about the possible health benefits from consuming each product. The researchers found that with the presence of a health claim, participants were more likely to rate the product effective in managing

health conditions not related to the product (for example, though sodium was not mentioned in any health claims on the labels, participants rated the foods as good for lowering blood pressure). Additionally, they rated the product as healthier and were more willing to pay; this supports the health halo effect.

In the discussion of their results, Roe *et al.* (1999) also explain that with a health claim (and less so with a specific nutrient claim), participants were less likely to explore the back of the package, which is generally where nutrition facts are located. Even if participants did decide to investigate the nutrition claims, if the health claim was on the front, participants were still more likely to state that they would purchase the product. Roe *et al.* (1999) used familiar brand-name packages with the brand labels removed; this may have created bias. If the participant already knew and trusted the brand, they may not have needed to check the back nutrition labels, and they may have bought it based on the front health claim alone. Furthermore, the may not even notice a health claim if they are in the habit of buying the same food; they may simply use the front label to check taste preference cues, like the size of curds in cottage cheese or flavor of fruit at the bottom of a yogurt cup.

Though this paper was written in 1999, before processed foods were as ubiquitous, its findings are still relevant today. The Food and Drug Administration's (FDA) current regulations stipulate that while a product may be low enough in fat to earn a "healthy" health claim label, it may be high in sugar, which is not currently mentioned in the FDA's daily value guidelines (which are listed next to the micro- and macronutrients). The FDA closely regulates what words may be used on food labels to avoid this confusion (U.S. Food & Drug Administration). Labels must adhere to the strict

guidelines regarding nutrient content claims, health claims, qualified health claims, and structure/function claims. However, the actual guidelines are more relaxed than most consumers realize. For example, in order for a package of food to bear the label "healthy," it must meet just one of the many possible requirements (be low in sodium, be low in fat, have a high DV of fiber) (Appendix 7.1). Likewise, a food that meets the requirement for low sodium may bear a health claim even though it may have a higher fat content. While the FDA protects consumers from false claims, some believe the lax regulations of the words "healthy" and "natural" should be changed (Balentine 2016). In response to this, the FDA has issued a formal request for comments from American citizens about their thoughts on "healthy" and "natural" labeling (Appendix 7.2)

For example, Kellog's Raisin Bran may be regarded as a "healthy" cereal—perhaps the word "bran" appeals to consumers wishing to increase their fiber intake (Kellogg's 2016). A box of Raisin Bran lists that it is "heart healthy," "made with REAL FRUIT," and is "perfectly balanced." While it is true that Raisin Bran has a high serving of fiber (7 grams, or 28% of the daily value of fiber), it also offers 18 grams of sugar—this is 50% of the recommended daily intake of added sugar for men, and 72% of the recommended daily intake of sugar for women (American Heart Association 2017). If a consumer consciously chooses to start her day with a "healthy" cereal, she may feel this choice licenses a bad choice later in the day. However, if she did not read the nutrition facts, as the results from Roe *et al.* (1999) suggest, then she may not realize the actual toll on her health these "healthy" choices are taking.

It is common, as Nelson (1970) notes, for advertisers exploit information to lure consumers. Advertisers must be persuasive in their language—especially when

describing experience goods, like packaged food products. They must convince consumers to buy the good, possibly by describing it ways that may sound like scientific claims; they may rely on the word "healthy" to entice consumers. In his 1970 paper, Nelson finds that the indirect advertisement of a healthy lifestyle or health gains impact experience goods. While the value of search goods is straightforward, more persuasive advertising and labels are necessary for experience goods—a 10 gallon bucket requires much less packaging and advertising than a box of breakfast cereal.

Similarly, Tal and Wansink (2014) find that simply adding graphs and formulas to market experience goods increase consumers' beliefs about their efficacy. While the FDA mandates that labeling must not be false or misleading, they do not have clearly defined limits on adding scientific graphs and charts to food labels. This addition of a scientific label suggests to consumers that the product has scientific or medical support, and is therefore more trustworthy. In their 2014 study, Tal and Wansink (2014) present participants with two descriptions of pharmaceutical drugs. One uses only words to describe the drug's efficacy, while the other uses a combination of graphs and words. The same information is conveyed in both conditions, and the graphs display simple relationships (Appendix 7.3). Public health research recommends that signage and other educational materials be kept around an 8th grade literacy level; a paragraph may seem daunting, while a graph is more accessible (Badarudeen and Sabharwal 2010). While graphs and charts on labels are not false or misleading to educated consumers (per the FDA regulations), they may convey a message of science-backed health claims that may deceive the majority of 8th grade literate Americans. Advertisers can take advantage of this confusion.

2.4 Contributions to Existing Literature

While all of the papers reviewed explore aspects of either moral balancing (as in Khan and Dhar 2006, Khan and Dhar 2007, Brañas-Garza *et al*. 2013), or consumers' perception of health words (as in Roe *et al*.1999, Lindsay 2010, Tal and Wansink 2014), there is a gap in the existing literature combining the effects of health advertising and moral balancing and licensing patterns. From Khan and Dhar (2007), it is clear that people choose vice foods before virtue foods if they know they will be able to make the same choice the following week. While this study has moral balancing and licensing implications, it focuses on better understanding selfishness and self-control. Based on this study, it is not realistic to conclude people would make all food or activity choices in an alternating 1:1 vice to virtue ratio throughout the day. I included five different time slots so moral balancing and licensing in the moral sphere of health and wellness can be better examined. I also observed the influence of habit on daily decisions, as well as differing tastes and preferences.

Consequently, my study is at the intersection of behavioral economic and public health. I attempt to address the question of whether or not the subjective FDA-regulated word "healthy" has a moral balancing or licensing effect on consumers. From on my findings, I explore avenues for further study and policy change in the conclusion section (Section 6).

2.5 Hypotheses

Based on the results of the sequential dictator games in Brañas-Garza *et al.* (2013), I expect to observe similar results. If moral balancing and licensing were present, that would provide evidence that the word "healthy" alone on a package, regardless of its

true health benefits, does have the power to influence consumers' decisions. However, if there is not licensing and balancing pattern, this may be because the participants understood the instructions that state that there is no difference in nutritional quality between food types. This could have created biased results; however, I hope to mitigate some of the effects of this by using a large sample size. I also included neutral activities (chess and reading). While these activities may be perceived as "healthy," they are not in the moral space of physical health and physical wellness. Based on Khan and Dhar (2007), I hypothesize these activities will not have an effect on how participants balance or license future decisions.

In addition to examining the data for signs of individuals' balancing and licensing patterns, I also examine the effects of the categorical variables. By running separate regressions for each categorical variable of interest, I capture some cross sectional variations. I also assess different groups of participants within my sample through ANOVA testing. I hypothesize that participants who start with an unhealthy choice (StartWith = 1) would have a lower average DailyScore than participants who started with neutral or healthy choices (StartWith = 2 or 3); participants in the latter categories would have net DailyScores closer to zero. If that is true, it would indicate that starting with a healthy or neutral choice has a beneficial impact on the participant's day.

3 Methodology and Data

3.1 Instrumentation

Because this study is new to the literature, there was no existing data; I created a short survey to answer my research question. It consisted of two parts: one drag-and-drop section to address moral balancing, and a short demographic survey. The full survey can

be found in (Appendix 7.7). To determine whether participants viewed "healthy" as a morally balancing or licensing word, pictures of thirteen different food and activity options were offered. Participants dragged and dropped their chosen food or activity into one of the five timeslots during their hypothetical day. Then, the times of day were coded, as were their choices. For example, a code of $1 = 1$, followed by $2 = -3$, and $3 = 2$, means that the participant chose an unhealthy food or activity at 8am, a healthy food or activity at 10am, and a neutral activity at 12 noon (no neutral foods were offered).

These thirteen options for choice and meal were chosen arbitrarily, based on availability of stock photos for Photoshop and a general consensus of what constitutes a healthy or unhealthy food. Basic healthy foods are outlined by public health organizations, and information about healthy eating is briefly taught in public schools; it follows that the population would generally agree about which foods are healthy and which are unhealthy. However, health preferences are unique to each individual; it would be difficult to understand how people engage in dietary and lifestyle moral balancing without asking participants outright. In Khan and Dhar (2007), participants choose between two foods—yogurt and a cookie. These choices are also arbitrary, based on general views of unhealthy and healthy foods. Future studies would benefit from pre-testing a variety of foods to determine a consensus on what foods constitute as healthy or unhealthy.

Personal taste may have skewed the results. In Khan and Dhar (2007), participants may have simply preferred the taste of a cookie to the taste of yogurt. Nutrition and health may not have been strong motivators for choosing the yogurt. Furthermore, some people may associate healthy foods with unpleasant tastes. Nutritionists report that

American diets are lacking in dark leafy greens (McMillan 2015). While this is due to a multitude of factors, one important thing to note is that kale may not taste as good to some people as iceberg lettuce. Personal taste may have skew the results regarding activities. I offered seven different activities, but participants may have not been particularly drawn to any of them. Additionally, the laptop and TV remote may be seen as substitutes; this would not necessarily have biased the results since both are unhealthy and coded as unhealthy = -1 in the second set of regressions, but since streaming and television programs are available online, it is important to be aware of this when drawing conclusions based on activity. In the third set of regressions, I coded laptop as neutral, as some participants may have conceptualized the laptop as part of a work or school day; it would fall into a different moral sphere, and therefore not factor into the balancing or licensing decision.

For my survey, I manipulated photos of food in Adobe Photoshop to portray different packaging. For the food types (cereal and cheeseburger) there were two options: "healthy" or "delicious" as shown in Figure 3. Ice cream was either advertised as "natural" or "delicious."[5] This is similar to the manipulation in Roe *et al.* (1999); participants were shown mock-ups of food packaging designs with health claims. However, instead of using brand name packages with the brands removed, I constructed a simpler package, where reading the few lines of text was essential for understanding if

[5] In 2015, the FDA announced it would be changing the guidelines for foods to bear the "natural" label. Because "natural" is another word that carries a potentially misleading health halo, I included it in my study to describe the healthy ice cream choice. The main effect I am looking for relates to the word healthy; however, natural is also a subjective FDA-regulated word undergoing review. I decided to include it in my study, and I will analyze the choice of "natural" ice cream as a "healthy" choice (U.S. Food and Drug Administration 2016).

the product was healthy or unhealthy. Pictures for unhealthy and healthy food options were identical.

Some activities and foods were "healthy" (playing basketball, gardening, running on the treadmill; foods were labeled "healthy"), and some were unhealthy (watching TV, working on a laptop; foods were not specifically labeled "healthy," and instead were called "delicious"). However, the description before the drag-and-drop question told participants that all meals were nutritionally equivalent. Though a food item said "healthy" on it, there was no nutritional difference from the alternative that said "delicious" (Appendix 7.4). If moral balancing and licensing occurrs, this would indicate that the presence of the word "healthy," disregarding its actual health benefits, has the power to license or balance decisions. However, because the statement about equivalent nutritional value is contrary to generally accepted notions of healthy and unhealthy food, picking a cheeseburger at all may have been unnatural to some participants; this could have create biased results. I also included neutral activities (chess and reading, in the second set of regressions; I added computer time to neutral activities in the third set) to fill out the day. It may have been difficult for participants to plan their day solely based on a schedule of eating and active or sedentary activities in the moral space of physical health and wellness. Because they are in a different moral space (playing chess may not make someone feel physically healthier or unhealthier) these are neutral activities that likely did not factor into the moral balancing and licensing pattern (as per the conclusions in Khan and Dhar 2007).

Five different time slots were shown (8am, 10am, 12 noon, 3pm, and 7pm). Participants created their day by selecting from the photos of the activities or foods and

dragged them into the time slot boxes. Participants could only perform one activity or eat one food at a time (multiple pictures per one time slot was not allowed), though they could choose to opt out of any answers (leaving boxes blank was acceptable). The next section was the demographic survey. Age, gender identity, location, and income data was collected. Participants were allowed to manually enter their age and US state they currently live in. Location is important to the survey not only because it may explain regional differences and social norms in health, but also because I wanted participants to have the same familiarity with and exposure to American food advertising; I sorted location narrowly, by Census Bureau divisions, and more broadly, by Census Bureau regions (U.S. Census Bureau 2015) (Appendix 7.5). To collect data on gender, which is now regarded as having a fluid, non-binary response, the question was posed as such: What is your gender identity? Possible responses were: Female, female to male transgender, male, male to female transgender, not sure, other, I prefer not to say. I focused on gender identity rather than sex, because gender identity is more related to my study than sex at birth (GenIUSS Group 2014). Income brackets were chosen based on American Community Survey (ACS) (U.S. Census Bureau 2015). I will discuss trends in the data and possible sources of selection bias in that data section (Section 3.3).

3.2 Procedure

The survey was created using the online academic survey site Qualtrics, and was administered using Amazon Mechanical Turk (mTurk). I completed Institutional Review Board (IRB) training and obtained IRB approval before the survey was published for responses.

Amazon Mechanical Turk is a marketplace for researchers. It offers Human Intelligence Tasks (HITs), or surveys, to workers on the mTurk cite. Anyone with an Amazon.com account can become a worker by registering; this involves a mock HIT that workers might encounter. Some HITs are long surveys, sometimes spanning the course of several days—there may be higher qualifications a worker must meet to accept this task. Additionally, as a reward for good work on the site, researchers can rate workers after they are finished completing the survey. Because some surveys offer monetary incentives, a researcher might rate a worker highly if the worker took time to thoughtfully answer questions (Amazon Mechanical Turk 2017). The process of becoming a worker on mTurk could be difficult for people who are elderly or unfamiliar with technology, and the time cost of setting up a worker account may be unappealing to people who are satisfied with their income. This creates a biased mTurk population from which to sample.

300 unique Human Intelligence Tasks (HITs) were offered (n=300). Workers were allowed 45 minutes to complete the survey, though the description advertised the survey would take between 3-5 minutes. A $1.00 monetary incentive was awarded to workers, which they received after 3 days. The sample size of 300 is appropriate for this survey. In previous studies on moral balancing that use undergraduate students as participants, the number of participants rages from 60 people to 180 people (Khan and Dhar 2007; Brañas-Garza *et al.* 2013). However, those studies were completed in person at a testing site. Because this survey was administered online, I could reach a larger population.

After reading the title ("Answer a brief survey about planning your day"), workers read the description ("If you accept this task, you will be redirected to a brief survey that will take less than 5 minutes to complete."). The monetary compensation was also displayed at this time.[6] A payment of $1.00 is appropriate for a 5-minute survey, as it results in an hourly wage of $12.00. To discourage workers just looking for money, the additional requirement that workers have a 90% worker rating was added to the program.

3.3 Data

The participants surveyed were registered Amazon Mechanical Turk workers, over 18 years of age. Additional worker requirements were location (people in the United States), and HIT approval rate (greater than 90%). These qualifications may result in some selection bias. After I had collected the data, I assigned all of the categorical variables numbers (Appendix 7.6). The total number of participants in the survey was 300 people (n=300). Preliminary data was collected before participants were excluded from the sample because they did not answer all questions of the survey.

There were a total of 1,438 activities and foods chosen across the full sample (n=300); the totals for each category can be found in the table below (Table 1). The most frequent choice was time on the computer; this may either signify work or schoolwork, to some participants, or leisure time. If computer time can be contextualized in this different moral space than health and wellness, it may be counted as a neutral, rather than unhealthy choice. I explore both of these ideas in the results section (Section 4). After time on the computer, the most frequent activity choice was running on the treadmill,

[6] A monetary incentive was given to participants to complete the survey. I was granted $210.00 by the Student Opportunity Fund at Skidmore College for payment to participants. The total fee for using mTurk was $420.00, and Professor Goff and I are working together to raise the other half of the funds. mTurk charges an additional 40% fee for participants after n=10 and for using their service.

followed by another neutral activity (reading), which suggests that participants may be trying to make healthy decisions. However, only one of the seven activity options was unhealthy; people who did not want to watch TV perhaps because they do not have a TV or cable may have avoided this option, instead choosing the laptop to stream content. For all the meal options, except for breakfast, more participants chose the healthy option than the unhealthy option. Cereal may be an outlier in this case because the unhealthy cereal was the first option at the top of the screen—participants may have chosen it without scrolling to see the other options.

Table 1: Frequency of participant food and activity choices (n=300)

CHOICE:	TOTALS
Healthy cereal	103
Unhealthy cereal	185
Healthy cheeseburger	193
Unhealthy cheeseburger	101
Healthy ice cream	47
Unhealthy ice cream	29
Laptop (Computer time)	243
Treadmill	142
Reading	138
TV Remote (Watching TV)	120
Gardening	97
Basketball	54
Chess	40

Seven participants were excluded from the initial sample, making the total sample population 293 people (n=293). Six of the seven participants were eliminated because they failed to make a choice for all five time slots; most of these participants skipped breakfast. One participant was removed because of failure to list a location where he or she currently lived in the US. Two participants who identified as "not sure" and "male to female transgender" are included in the "male and non-female" category (male and any non-female = 0). In the first regression, I further narrowed the sample size by excluding reading and chess as neutral activities—computer time was categorized as unhealthy. The smaller sample includes 125 participants (n=125). The descriptive statistics can be found in the table below (Table 2).

Table 2: Descriptive statistics for categorical variables (n=125)

N=125	VARIABLE (HIGHEST % FREQUENCY)	VARIABLE (LOWEST % FREQUENCY)	MEAN
Age	19 years (0.80)	75 years (0.80)	34.7 years
Gender	Male or any non-female (64.00)	Female (36.00)	N/A
Region	South (36.00)	Northeast (17.60)	N/A
Income	$50,000-$74,000 (23.21)	$200,000 or more (0.68)	Between $25,000-$50,000

In the next set of regressions I ran, I used 293 participants, and I coded on a scale: unhealthy = -1, neutral = 0, healthy = 1. Chess and reading were categorized as neutral (0), and time on the computer was categorized as unhealthy (-1) The descriptive statistics for the larger sample can be found in Table 3 on the next page.

Table 3: Descriptive statistics for categorical variables (n=293)

N=293	VARIABLE (HIGHEST % FREQUENCY)	VARIABLE (LOWEST % FREQUENCY)	MEAN
Age	19 years (0.34)	75 years (0.34)	36.0 years
Gender	Male or any non-female (60.07)	Female (39.93)	N/A
Region	South (34.25)	Northeast (19.86) and Midwest (19.86)	N/A
Income	$50,000-$74,000 (23.21)	$200,000 or more (0.68)	Between $25,000-$50,000

The most frequent daily score for this group was a neutral score of 0, with 80 participants having a net daily score of zero (Table 4). The next most frequent scores were -1, with a frequency of 62 participants, and a score of +1, with 59 participants. Though this does not directly indicate balancing or licensing necessarily, it shows that participants may have been aware of the healthy and unhealthy effects their choices had on their days. Another interesting point to note is that most participants (191 participants) started with an unhealthy choice at 8am. Only 100 participants began with a healthy choice, and six participants began with a neutral choice. I discuss the effect of the starting choice on the rest of the day in the results section (Section 4).

In the final set of regressions I ran on the larger sample (n=293), I categorized time on the computer as neutral. The descriptive statistics are the same as in the table above (Table 3). I also added another category based on region—this groups locations by larger geographical region (West, Midwest, South, Northeast), as defined by the U.S. Census Bureau. An overwhelming number of participants were from the South (100); from the West there were 77 participants, and from both the Midwest and Northeast there were 58 participants. However, the high number of participants from the South is likely because the South is comprised of many more states than the other regions (see Appendix

7.5). In this construction of the data, where computer time was neutral, no participants had a DailyScore of lower than -3 (and only 4 participants scored -3).

Table 4: Frequency of Daily Scores counting computer time as unhealthy vs. neutral

DailyScore	Freq. computer time = unhealthy	Freq. computer time = neutral
-4	3	0
-3	12	4
-2	40	19
-1	62	64
0	80	71
1	59	76
2	33	42
3	6	15
4	2	2
Total	293	293

The mode income across the samples was consistent with the 2015 median household income ($55,775) (Posey 2016). However, the sample may not be fully representative of the population of U.S. consumers. This is due to the limitations of using mTurk and the formatting problems I encountered by issuing the survey online. mTurk only reaches people with Amazon.com accounts who register to be workers for the site. Additionally, because of monetary and academic research incentives, the population may include a proportionally higher number of students. Navigating to the site and learning how to use it may pose challenges for older individuals who are less familiar with technology. This could explain the low median age (34.6); the mean age of the U.S. population in 2016 was 37.9 years (Central Intelligence Agency). Additionally, because mTurk is used in academic studies, researchers in undergraduate and graduate schools may be more familiar with the site and be more willing to participate, hoping that other workers will take their surveys.

3.4 Empirical Specifications

To analyze the panel data, I ran preliminary fixed-effects regressions. Later, I will run random-effects regressions. I am using the model that Brañas-Garza *et al.* (2013) used to explore their data.

$$d_{it} = \alpha_i + \gamma d_{i,t-1} + x_{it}\beta + \varepsilon_{i_t}$$

In this equation, *d* is the dummy variable for choice. In the initial regression, I coded "unhealthy" as 0 and "healthy" as 1. In the second set of regressions, I assigned choices on a scale from -1 to 1, with "unhealthy" equal to -1, "neutral" equal to 0, and "healthy" equal to +1. For the final draft, I will calculate participants' net daily scores and run an independent t-test to determine which strategy (starting with a healthy or unhealthy choice) yields the most balanced day (with a net score close to 0). The explanatory variables ($x_{it}ß$) I use include age, gender identity, location in the U.S. based on IP address, and income. I coded these categorical variables in Appendix 7.6.

While I use the equation in Brañas-Garza *et al*. (2013), I will not be using GMM estimators to evaluate the data, instead I use OLS estimates. I realize this is not ideal for my data set, because I have a large sample size but a small number of time slots (known as the "small T, large N" problem) (Ge and Ho 2017). Additionally, I lag one of the variables (decisions after $t-1$). This creates a problem with analysis of the panel data. OLS estimates are more suited for data over a longer period of time (i.e. with a "large T") and will bias my results towards the upper bound. However, I also evaluate fixed effects, which estimates the lower bound.

4 Results

The panel regression can be found in Table 4 below. For this regression, I used the smaller dataset that excludes neutral (chess and reading) choices. The negative coefficient indicates that the previous period's choice (period 1) will lead to a 17% chance that the next period (period 2) will not be different. Therefore, based on the previous choice, there is a 17% chance that the next chance will be the same. This implies strong moral balancing effects between the first and second periods. These results are consistent with the results found in Brañas-Garza *et al.* (2013). However, these results do not persist in the second lag.

Table 5: Effects of balancing and licensing at Lag.1 (n=125), no neutral activities considered

Dependent Variable: choice Constant: 0.624	Coefficient (Standard Error)
Lag1.Choice	-0.169*** (0.052)

*** p<0.01

In the second regression, I used the entire dataset (n=293) and coded unhealthy, neutral, and healthy choices on a scale from -1 to +1. Computer time is included as an unhealthy choice in this version of the model. With the larger sample, I found similar balancing and licensing patterns as in the smaller dataset. Next, I investigated if moral balancing patterns were consistent if computer time was coded as a neutral (computer time = 0) activity. For all these preliminary regressions, using the smaller sample, large sample with computer time coded as unhealthy, and large sample with computer time coded as neutral, moral balancing and licensing patterns were strong. There was between a 14.5% chance (when the laptop was coded as neutral) and 16.5% chance (from the smaller sample) that the first two choices would be the same.

However, I found that at lag2, for all samples (small sample excluding neutral activities, large sample with laptop coded as neutral and as unhealthy) the previous choice was around 40% likely to not be different. Table 6 displays the data for the regression that considers lags 1 and 2 when the laptop was coded as unhealthy. This demonstrates that there is a 40% chance that the first choice and the third choice will be the same. While the moral balancing and licensing is not as strong at lag2, the effect is still statistically significant. The diminution of this effect may be due to habit. Another possibility is that it may indicate a longer balancing pattern; participants might not make balancing decisions in an alternating 1:1 healthy:unhealthy ratio, but perhaps one healthy choice licenses two or more unhealthy choices. Brañas-Garza *et al.* (2013) found similar patterns; though balancing did not occur strictly in an alternating 1:1 ratio, the donation amount tended to even out over time. I observe the same effect in Table 3—participants tried to keep their net daily scores close to zero.

Table 6: Effects of balancing and licensing at Lag.2 (n=293), laptop coded as unhealthy

Dependent Variable: Choice Constant: 0.209	Coefficient (Standard Error)
Lag1.Choice	-0.395*** (0.032)
Lag2.Choice	-0.401*** (0.031)

***p<0.01

I also ran individual regressions focusing on the effect of the categorical variables. By running each regression individually, I observed some of the random effects present in the model (shown in Table 7). The results in the table consider each of the variables individually; I combined the results in one table to make it easier to read; however, each variable was run in a separate regression (for example, the choice after one

lagged period if participants were between 18 and 44 years old was a separate regression).

Table 7: Individual regressions considering categorical variables (n=293), laptop coded as unhealthy

Dependent Variable: Choice	Coefficient (Standard Error)
L.Choice (if age 18-44)	-0.1539 (0.0325)***
L.Choice (if income $25,000-$34,999)	-0.1549 (0.0906)***
L.Choice (if income $100,000-$149,999)	-0.1208 (.1145)

*** p<0.01

These patterns show a similarly weak balancing and licensing pattern as in the original data. This is likely because most of the participants fell between the ages of 18 and 44, and the income bracket $25,000-$34,999 was the second most frequent income bracket in the data (50 responses).

Next, I ran ANOVAs on the full dataset (n=293) for computer time as both unhealthy and neutral (Table 8). I have only included the ANOVA results for categorical variables when computer time was coded as unhealthy because the results when computer time was coded as healthy had the same balancing and licensing implications. Because ANOVA analysis exposes statistically significant differences among the means across groups, I could analyze some of the categorical variables' effect on participants' overall daily scores. For example, I answer the question: "Does starting the day with a unhealthy, neutral, or healthy choice have a significant impact on the DailyScore?" I find that when time on the computer is coded as either unhealthy or neutral, the mean of the DailyScores for each type of starting score (unhealthy, neutral, or healthy) is significantly different. This indicates that there is a strategy for how to start the day that results in ending the day

with a net score close to zero. T-tests confirmed moral licensing and balancing effects. Though the other categorical variables were responsible for differentiation in DailyScores, they were not statistically significant, so I did not examine whether or not t-tests indicated moral balancing or licensing. This could be because health trends vary more widely across age, gender, income, and region than I had predicted.

Table 8: ANOVA tests on various categorical variables when computer time is unhealthy

Dependent Variable: DailyScore	Indication of Moral Balancing and Licensing Effect	P-value
Starting the Day With Unhealthy (-1), neutral (0), or healthy (1)	Moral balancing indicated	0.00
Age	Means are not statistically different; no indication of moral balancing and licensing	0.351
Gender	Means are not statistically different; no indication of moral balancing and licensing	0.887
Income	Means are not statistically different; no indication of moral balancing and licensing	0.550
Region	Means are not statistically different; no indication of moral balancing and licensing	0.681

5 Discussion

I did not observe the starkly contrasting moral balancing and licensing effects that were seen in Brañas-Garza *et. al* (2013). There are several possibilities for why this happened. First, I told the participants that all meals were nutritionally equivalent. Therefore, the presence of the word "healthy," might not have had any effect because in

this experiment, “healthy” was meaningless. Additionally, habit, personal taste preferences, and selection of participants may have led to biased results.

Based on Khan and Dhar (2007), I predicted that participants would make an unhealthy choice first. The data indicate that most people picked the unhealthy cereal to start their day, but this could be biased, as the unhealthy cereal choice was at the top of the overall choice list. They may have seen cereal and picked it, completely disregarding the packaging. Having only a meat option for lunch may have created bias; vegetarians may have picked either the healthy or unhealthy cheeseburger randomly, since the nutritional claim would not have affected them.

My survey design may have also created biased results. The drag and drop method may have been difficult to use, as it worked best with Google Chrome and had some technical display issues on Safari. I tried to ameliorate this by telling participants before starting the survey that Chrome was the preferred browser. Another point of bias in the survey design is in the choices I allowed participants to pick from. Though I offered an equal number of healthy and unhealthy foods, I offered only two unhealthy activities (or only one unhealthy activity, if time on the computer is counted as neutral) and three healthy activities. Having more options available for healthy choices skews the results towards healthy activity, leading to an overestimation of participants’ interest in health and wellness. However, the most frequent activity after the computer time was the treadmill, followed by reading (a neutral activity). As discussed previously, watching TV may have not been chosen because laptops and TVs are substitutes. An additional explanation is that the participants chose healthier options because the survey was to self-

report a planned activity. Khan and Dhar (2007) illustrates that people are optimistic about their future choices; this may have skewed the results towards healthier decisions.

5.1 Policy Implications

In my survey, participants read "healthy" to truly be more beneficial than the packages that did not say "healthy," even though they were explicitly told that the nutrition facts are equivalent. This led them to choose other foods and activities in their day in a moral balancing and licensing pattern; the health halo around the word "healthy" is strong enough to elicit balancing and licensing patterns. In the survey, the participants chose a "healthy" cheeseburger, and some participants followed this decision with an unhealthy activity (watching TV). Though they thought they were making a healthy food choice that licensed the unhealthy activity, they were actually making two unhealthy choices. Unfortunately, this could contribute to a growing unhealthy population because consumers may make balancing and licensing decisions in the presence of the word "healthy," even though it is not nutritionally superior.

Further studies on "healthy," "natural," and other health-related words that may have a health halo effect are necessary. Based on my study, I suggest that the FDA continue its reconsideration of the subjective health words "natural" and "healthy." Additionally, I believe they should study the effects of the words "real," "100%," and "whole," to see if they have the same effects on moral balancing and licensing. The FDA should reconsider the regulation of all subjective health words so that consumers can trust food packaging and labels and practice moral balancing and licensing without inadvertently suffering adverse health effects.

6 Conclusion

As stated earlier in the discussion section (Section 5), I did not find the clear results I had predicted; though there is some degree of moral licensing and balancing, the pattern is not as starkly displayed as in Brañas-Garza *et al.* (2013), which I partially modeled my study after. This could be due to many different reasons. First I will discuss issues with the current version of the survey that may have led to these results. Next, I will explore how the survey could be changed to answer related research questions based on the results of the survey.

6.1 Limitations of the current survey

Based on Brañas-Garza *et al*. (2013), I hypothesized I would see clear moral balancing and licensing effects. While I did find that after the first lag, there were significant moral balancing and licensing effects, after the second lag, this became less prominent. Due to my survey design and other limitations, I did not exactly follow their method; this likely yielded a less clear balancing and licensing pattern. First, my survey only had five time slots, whereas participants in Brañas-Garza *et al*. (2013) played sixteen sequential dictator games. This further compounds the "small T, big N" problem discussed in Ge and Ho (2017). With more time slots, it may be easier to observe clear balancing and licensing patterns, even if they do not follow a 1:1 healthy : unhealthy ratio.

However, five time slots may have presented too many options for people. For many people, especially in the 18-44 age range, weekdays follow a structure around work and school. This may also explain some issues with the 1:1 ratio of alternating vice;virtue in terms of habits. For example, if a participant does not eat breakfast in the morning, she

may have chosen “TV” for 8am and “Laptop” for 12 noon. This would result in a -1,-1 response. She may not conceptualize watching TV at 8am as an unhealthy choice; rather, it is her habit instead of eating breakfast. Offering more time slots would exacerbate this problem. Additionally, people may not plan for eating or doing five different things in one day (e.g. someone could get up, go to work, eat dinner, and go to bed). Furthermore, asking people who do not routinely plan their day to think about how they would structure their day likely creates some bias. While one suggestion could be to ask participants to design a week full of five-choice days, this may also be unrealistic for many people, as they may not plan a day or even a week with the same inclinations towards moral balancing and licensing.

Another potential issue in my survey was the lack of variety in foods and activities offered. With more time slots, participants could choose a more varied day. A pre-survey could be administered to a different group of participants, asking them about perceived healthiness and unhealthiness of certain food types. More realistic packages could be rendered in Photoshop, and the healthiness and unhealthiness of these could be rated as well, to get a baseline attitude towards the foods as healthy or unhealthy. I could also offer more variation between the healthy or unhealthy foods. For example, instead of offering a cereal that is explicitly healthy (healthy = 1), there could be healthy cereals that vary in degrees of healthiness (e.g. plain oatmeal = 80, Raisin Bran = 20). Ultimately, however, health preferences are specific to individuals, so a pre-survey group may not have the same perceptions of health as the survey group, leading to biased results. The order of the pictures was the same for each participant; the first picture shown was unhealthy cereal, which may have led participants to automatically choose it without

scrolling to see the other options. In future study, randomizing the order of the foods and activities would eliminate this bias.

I also may have encountered bias my survey due to the fact the pictures of the different foods labeled “healthy” or “delicious” may have looked too similar. It makes sense that I did not observe a balancing and licensing pattern if people were unaware of a change in packages. More planning and experience in Photoshop could ameliorate this problem if I were to run the experiment again in the future.

While food and exercise may be in the same moral sphere for some people, others may weigh them differently. Keeping the options in one moral sphere (either all food or all activities) may produce a clearer balancing and licensing pattern; however, it would be difficult to increase the number of time slots, as sixteen different episodes of activity or eating in one day is excessive and unrealistic.

Age-based selection bias may have also skewed the results. The mean age of participants was 34.6 years; these participants would still probably feel comfortable running on a treadmill and playing basketball, but perhaps these activities may have seemed too strenuous for older people in the sample. To combat this, I also included the healthy activity gardening. However, healthy activities are also subject to individual tastes and preferences, as well as regional differences. Some participants may not have chosen the treadmill option because it is warm enough where they live and they like to exercise outdoors. Or, perhaps it is too cold to garden where people live, so they may have not picked that healthy option. Calibrating the exact number of options, without overwhelming participants, and choosing what options to provide, is a major challenge in this field of study.

6.2 Further Study

I did not observe an alternating healthy:unhealthy ratio. It would be interesting to study how many virtuous decisions license a vice decision, and if the magnitude of the virtue can license several moderate vice decisions. Pre-screening would determine which virtues and vices are more serious than others (walking to work is a slight positive, while going to the gym has a bigger effect; or, eating a piece of chocolate is a minor offense, while eating a piece of cake is a larger offense). Because the pattern of moral licensing and balancing is not an alternating 1:1 healthy:unhealthy ratio, differences in magnitude may explain variation in the balancing effect.

Pre-screening could also be implemented to determine what other words might have this effect. While I chose to use "healthy" and "natural" because they are under review by the Food and Drug Administration, words like "100%," "real," and "whole" may have similar health halo effects. A future study could count the frequency of these health buzzwords on packages, and then a pre-test could determine how consumers perceive these words. I would expect claims that bear these buzzwords would have similar balancing and licensing effects.

The time of day may also have had an impact on how participants answered the survey. A different study could present participants with vignettes about a healthy activity; for example, "you go to the gym in the morning," or, "you have an intramural soccer game tonight." Participants would be asked to structure their day around this event, balancing or licensing on the condition of a set virtue time. This is similar to the experiment Khan and Dhar (2007) ran with the yogurt and the cookie, and I would expect that people would balance and license in a similar pattern. People chose vice first when

they knew the next week they would get yogurt; in this proposed study, participants may choose vices if they knew later in the day they would perform a virtuous activity. Similarly, another future design could present participants with an activity initially ("pretend you have just been to the gym") and then ask them to evaluate a healthy or unhealthy choice immediately afterwards ("please rate how much you would like to eat this pizza on a scale of 1-100). This is more similar to the Khan and Dhar (2007) study.

Another interesting aspect of study is the ability of quantitative information on nutrition labels to produce a balancing and licensing effect. Individuals consider many different factors when they are choosing what to buy from the grocery store, like advertising, nutrition, and taste. However, different consumers may consider these factors from either a quantitative or qualitative aspect, or both. A health claim on the front of the package (qualitative) may prompt a consumer to check the back of the package (quantitative). Or, as in Roe *et al.* (1999), a front-label claim of health may be sufficient for consumers, and would not warrant a back-label search for nutrition. A survey could be administered with identical front-labels but different nutrition labels. One difficult aspect of both my survey and a future survey like this is the need to keep the packaging or the nutrition constant. This is unrealistic, and may bias the results.

The most comprehensive study may not be a survey before participants balance and license events and foods in their days, but it could be administered after participants have lived a regular day. A survey could be given at night, right before participants go to sleep. It could ask about what the participants ate and did during that day. While these reflections would be subject to poor recall and self-reporting, they may provide the most accurate description of a participant's day. The following day, the foods the participant

referenced could be examined for health words, and the activities could be coded. However, participants who make their own breakfast from ingredients (an egg, piece of bread, and slice of bacon) instead of participants who buy their breakfasts (a packaged breakfast sandwich), may have to be excluded from the study. To fix the unnecessary cost of surveying outliers, participants could be given a list of acceptable foods to eat throughout the week (packaged food bearing varying quantitative and qualitative health information).

Based on the results of my study, I believe the FDA should change the regulations of the words "healthy," "natural," and perhaps other words that create the health halo effect so that consumers can trust these claims and use them as guidelines, since consumers use subjective health words to balance and license decisions.

7 Appendix

7.1: Table from the Food and Drug Administration's regulation of the word "healthy." U.S. Food & Drug Administration (2016).

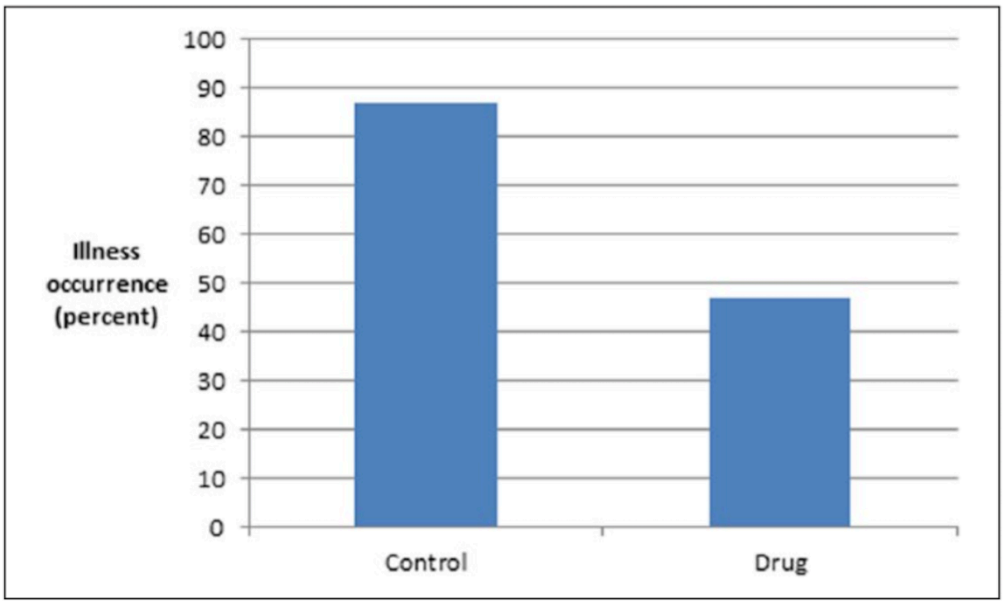
(i) The food meets the following conditions for fat, saturated fat, cholesterol, and other nutrients:

If the food is...	The fat level must be...	The saturated fat level must be...	The cholesterol level must be...	The food must contain...
(A) A raw fruit or vegetable	Low fat as defined in 101.62(b)(2)	Low saturated fat as defined in 101.62(c)(2)	The disclosure level for cholesterol specified in 101.13(h) or less	N/A
(B) A single-ingredient or a mixture of frozen or canned fruits and vegetables [1]	Low fat as defined in 101.62(b)(2)	Low saturated fat as defined in 101.62(c)(2)	The disclosure level for cholesterol specified in 101.13(h) or less	N/A
(C) An enriched cereal-grain product that conforms to a standard of identity in part 136, 137 or 139 of this chapter	Low fat as defined in 101.62(b)(2)	Low saturated fat as defined in 101.62(c)(2)	The disclosure level for cholesterol specified in 101.13(h) or less	N/A
(D) A raw, single-ingredient seafood or game meat	Less than 5 grams (g) total fat per RA [2] and per 100 g	Less than 2 g saturated fat per RA and per 100 g	Less than 95 mg cholesterol per RA and per 100 g	At least 10 percent of the RDI [3] or the DRV [4] per RA of one or more of vitamin A, vitamin C, calcium, iron, protein, or fiber
(E) A meal product as defined in 101.13(l) or a main dish product as defined in 101.13(m)	Low fat as defined in 101.62(b)(3)	Low saturated fat as defined in 101.62(c)(3)	90 mg or less cholesterol per LS [5]	At least 10 percent of the RDI or DRV per LS of two nutrients (for a main dish product) or of three nutrients (for a meal product) of: vitamin A, vitamin C, calcium, iron, protein, or fiber
(F) A food not specifically listed in this table	Low fat as defined in 101.62(b)(2)	Low saturated fat as defined in 101.62(c)(2)	The disclosure level for cholesterol specified in 101.13(h) or less	At least 10 percent of the RDI or the DRV per RA of one or more of vitamin A, vitamin C, calcium, iron, protein or fiber

7.2: FDA questions for public opinion on the word "natural"
U.S. Food & Drug Administration (2016)

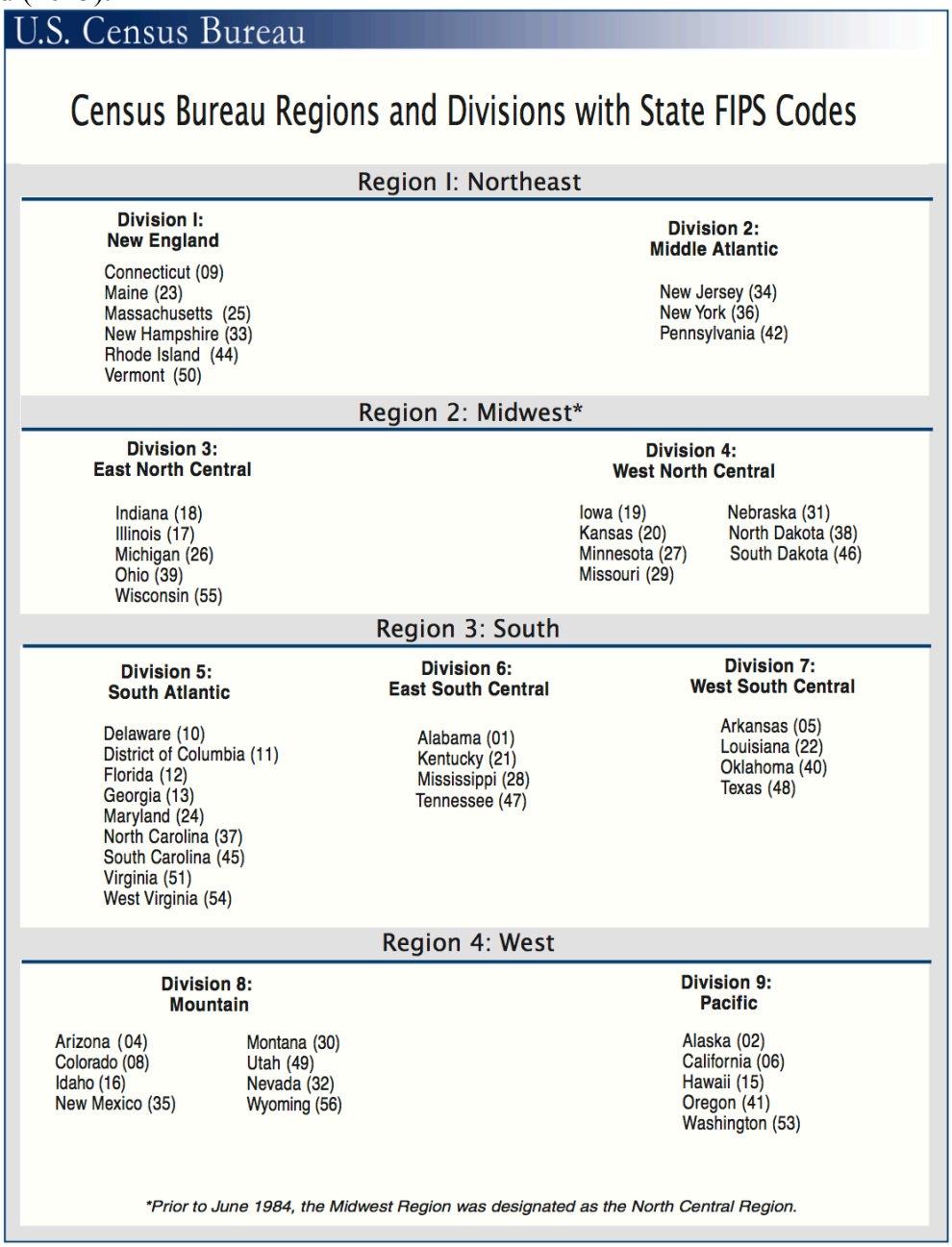
Whether it is appropriate to define the term "natural,"
If so, how the agency should define "natural," and
How the agency should determine appropriate use of the term on food labels.

7.3: Graph from Tal and Wansink (2013) that accompanied a short paragraph to explain the efficacy of the drug
Tal, A. and Wansink, B. (2014). Blinded with science: Trivial graphs and formulas increase ad persuasiveness and belief in product efficacy. *Public Understanding of Science*. 1-9.



7.4: “Healthy” and “delicious” hamburger options in the survey I designed

7.5: Census Regions and Divisions of the United States
U.S. Department of Commerce Economics and Statistics Administration U.S. Census Bureau (2015).



U.S. Census Bureau

Census Bureau Regions and Divisions with State FIPS Codes

Region I: Northeast

Division I: New England

Connecticut (09)
Maine (23)
Massachusetts (25)
New Hampshire (33)
Rhode Island (44)
Vermont (50)

Division 2: Middle Atlantic

New Jersey (34)
New York (36)
Pennsylvania (42)

Region 2: Midwest*

Division 3: East North Central

Indiana (18)
Illinois (17)
Michigan (26)
Ohio (39)
Wisconsin (55)

Division 4: West North Central

Iowa (19)
Kansas (20)
Minnesota (27)
Missouri (29)
Nebraska (31)
North Dakota (38)
South Dakota (46)

Region 3: South

Division 5: South Atlantic

Delaware (10)
District of Columbia (11)
Florida (12)
Georgia (13)
Maryland (24)
North Carolina (37)
South Carolina (45)
Virginia (51)
West Virginia (54)

Division 6: East South Central

Alabama (01)
Kentucky (21)
Mississippi (28)
Tennessee (47)

Division 7: West South Central

Arkansas (05)
Louisiana (22)
Oklahoma (40)
Texas (48)

Region 4: West

Division 8: Mountain

Arizona (04)
Colorado (08)
Idaho (16)
New Mexico (35)
Montana (30)
Utah (49)
Nevada (32)
Wyoming (56)

Division 9: Pacific

Alaska (02)
California (06)
Hawaii (15)
Oregon (41)
Washington (53)

**Prior to June 1984, the Midwest Region was designated as the North Central Region.*

7.6: Table of translated categorical variables in the survey to dummy variables

CHOICE (IN ORDER OF APPEARANCE IN THE SURVEY)	CODE	GENDER	CODE	LOCATION	CODE	INCOME	CODE
Delicious cereal	-1	Female	1	New England (District 1, Northeast)	1	Less than $10,000	0
Laptop (Study 3)	-1 (0)	Male or any other non-female	0	Middle Atlantic (District 2, Northeast)	2	$10,000-$14,999	1
Chess	0			East North Central (District 3, Midwest)	3	$15,000-$24,999	2
Natural ice cream	+1			West North Central (District 4, Midwest)	4	$25,000-$34,999	3
Healthy cheeseburger	+1			South Atlantic (District 5, South)	5	$35,000-$49,999	4
TV remote (watch TV)	-1			East South Central (District 6, South)	6	$50,000-$74,999	5
Run on the treadmill	+1			West South Central (District 7, South)	7	$75,000-$99,999	6
Healthy cereal	+1			Mountain (District 8, West)	8	$100,000-$149,999	7
Gardening	+1			Pacific (District 9, West)	9	$150,000-$199,999	8
Delicious ice cream	-1					$200,000 or more	9
Read	0						
Delicious cheeseburger	-1						
Basketball hoop (play basketball)	+1						

7.7: Qualtrics survey, administered in Amazon Mechanical Turk

Consent

1. INTRODUCTION
You are invited to be a participant in a research study being conducted by Sophie Tate, a student in the economics department at Skidmore College, under the direction of Dr. Sandra Goff, Assistant Professor of Economics at Skidmore College. We ask that you read this document and ask any questions you may have before agreeing to be in the study.

2. BACKGROUND
The purpose of this study is to investigate how individuals decide to structure their time and daily activities.

3. DURATION
Participation in this study is expected to take approximately 3-5 minutes.

4. PROCEDURES
If you agree to participate in this study, you will be asked to answer questions about how you might structure a hypothetical day. Then you will be asked to answer a few questions about yourself. Please consider the questions carefully. You may choose to skip any of the questions. You may also quit at any time, however you must complete the study to receive compensation.

Please note that you may participate in this study only once. If you choose to quit this study at any time, you will NOT be eligible to accept the task later.

5. RISKS/BENEFITS
There are no foreseeable significant risks of participation in this study as the study requires participants to make choices typical to daily life.

The benefit of participation is that it contributes to research regarding consumer and lifestyle choice.

6. CONFIDENTIALITY
The records of this study will be kept private. In any sort of report that is published or presentation that is given, we will not include any information that will make it possible to identify a participant.
Data will not be saved with identifiers. Only the researchers will have access to the data, which will be kept in a password protected file on the researchers' computers. IP addresses will not be recorded.
Please recognize that we are using an online survey to collect this data, and have taken all reasonable measures to protect your identity and responses. For example, the data is SSL encrypted, it is stored on a password protected database, and IP addresses are not collected. These measures provide the high level of security that is used by financial institutions and it is very unlikely that your data could be accessed by anyone. However, e-mail and the Internet are not 100% secure. Therefore, we also suggest that you clear the computer's cache and browser history to protect your privacy after completing the survey.

7. COMPENSATION
If you meet certain requirements, you will receive $1 for completing the study. These requirements include:

- being 18 years of age or older
- being able to read and understand English
- residing in the United States of America
- reading all instructions and completing the survey in good faith
- answering questions correctly which are designed to determine if you read and understood all instructions and survey content
- finishing the survey and receiving a confirmation code

This follows Amazon Mechanical Turk policy, which states that "a Requester may reject your work if the HIT was not completed correctly or the instructions were not followed."

Participants cannot be compensated without a valid confirmation code.

8. VOLUNTARY NATURE OF THE STUDY
Your decision whether or not to participate is entirely voluntary. You may refuse to participate before the study begins, discontinue at any time, or skip any questions and/or procedures that make you feel uncomfortable, with no penalty to you, and no effect on your academic standing, record, and/or current or future relationship with Skidmore College or any of its representatives.

9. CONTACTS AND QUESTIONS

The main researcher conducting this study is Sophia Tate, an undergraduate student in the economics department at Skidmore College. If you have any questions, now or later, **you may contact Sophia Tate at state@skidmore.edu or Ms. Tate's advisor, Dr. Sandra Goff, at sgoff@skidmore.edu or (518) 580-5095**.

If you have questions or concerns regarding this study and would like to speak with someone other than the researcher, **you may contact Mary Hoehn, Institutional Review Board Chair, Skidmore College, at (518) 580-8052 or mhoehn@skidmore.edu**.

10. STATEMENT OF CONSENT

You may print a copy of this form for your records.

The procedures of this study have been explained to me and my questions have been addressed. The information that I provide is confidential and will be used for research purposes only. I understand that my participation is voluntary and that I may withdraw anytime. If I have any concerns about my experience in this study (e.g., that I was treated unfairly or felt threatened), I may contact the Chair of the Institutional Review Board or the Chair of the sponsoring department of this research regarding my concerns.

If you agree to take the survey, please indicate your consent by clicking "I have read the above information and agree to participate." If you decline participation, please indicate by clicking "Please remove me from this study." Continuing with this study implies your consent to participate.

By clicking the circle below at left you indicate that you have read the above information and agree to participate.

I have read the above information and agree to participate	Please remove me from this study.
○	○

Are you sure you would like to be removed from this study? If you choose to quit, you will not receive compensation

Yes, I would like to quit.	**No, I would like to continue the study.**
○	○

Choices

These page timer metrics will not be displayed to the recipient.

First Click: *0 seconds*

Last Click: *0 seconds*

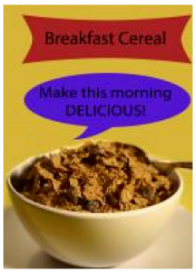
Page Submit: *0 seconds*

Click Count: *0 clicks*

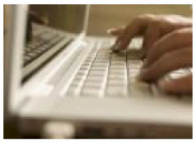
Drag and drop the pictures on the left into the boxes to organize your ideal day. Assume all meals/foods of a similar type are nutritionally equivalent, meaning that they contain the same amount of fat, protein, carbohydrate, sugar, vitamins, minerals, calories, etc.

Before making your selections, please make sure to carefully review all possible choices. This may require scrolling.

8am



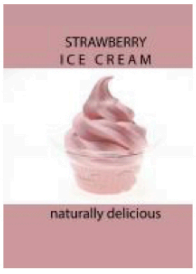
10am

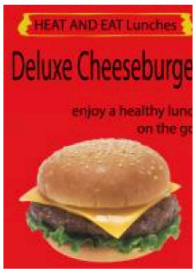


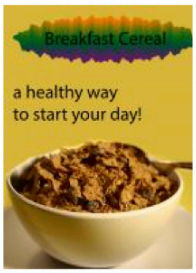
12 noon

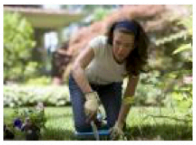
3pm

7pm






Breakfast Cereal
a healthy way
to start your day!




cold, sweet, delicious
ICE CREAM

Block 2

These page timer metrics will not be displayed to the recipient.
First Click: *0 seconds*
Last Click: *0 seconds*
Page Submit: *0 seconds*
Click Count: *0 clicks*

What is your age, in years?

What is your gender identity?

In which US state do you currently live?

What is your household income in the past 12 months?

- Less than $10,000
- $10,000-$14,999
- $15,000-$24,999
- $25,000-$34,999
- $35,000-$49,999
- $50,000-$74,999
- $75,000-$99,999
- $100,000-$149,999
- $150,000-$199,999
- $200,000 or more

References:

Amazon Mechanical Turk (2017). *General Questions*. Retrieved from: https://www.mturk.com/mturk/help?helpPage=overview. Accessed April 6, 2017.

American Heart Association (2017). Added Sugars. Retrieved from: http://www.heart.org/HEARTORG/HealthyLiving/HealthyEating/Nutrition/Added-Sugars_UCM_305858_Article.jsp#.WQkQ0hiZPdQ Accessed March 20, 2017.

Balentine, D. (2016, September 27). Making Sure 'Healthy' Means What It Says on Food Packages [Web log post]. Retrieved from https://blogs.fda.gov/fdavoice/index.php/2016/09/making-sure-healthy-means-what-it-says-on-food-packages/. Accessed February 17, 2017. x

Badarudeen, S., and Sabharwal, S. (2010). Assessing Readibility of Patient Education Materials: Current Role in Orthopaedics. *Clinical Orthopaedics and Related Research* 486(10). 2572-2580.

Brañas-Garza P. *et al.*. (2013). Moral Cleansing and Moral Licenses: Experimental Evidence. *Economics and Philosophy*, *29*, 199-212.

Cason, T.N., and Mui, V-L., (1998). Social Influence in the Sequential Dictator Game. *Journal of Mathematical Psychology*, *42*, 248-265.

Central Intelligence Agency. The World Factbook: Median Age. Retrieved from: https://www.cia.gov/library/publications/the-world-factbook/fields/2177.html . Accessed April 3, 2017.

Dahl, D. W., *et al*. (2003). The Nature of Self-Reported Guilt in Consumption Contexts. *Marketing Letters, 14*(3), 159-171.

Delaney, M., and McCarthy, M.B. (2013). Saints, Sinner, and Non-believers: The Moral Space of Food. A Qualitative Exploration of Beliefs and Perspectives on Healthy Eating of Irish Adults Aged 50-70. *Appetite, 73*, 105-113.

Ge, Q and Ho, B. (2017). Energy Consumption and Habit Formation: Evidence from High Frequency Thermostat Usage Data. Working paper. 1-52.

GenIUSS Group. (2014). Best Practices for Asking Questions to Identify Transgender and Other Gender Minority Respondents on Population-based Surveys. J.L. Herman (ed.).

Jallinoja, P. *et al.* (2010). Negotiated Pleasures in Health-Seeking Lifestyles of Participants of a Health Promoting Intervention. *Health, 14*(2), 115-130.

Kellogg's Raisin Bran Cereal. (2016). Retreived from: https://www.kelloggs.com/en_US/products/kellogg-s-raisin-bran-cereal-product.html#nutrition-modal. Accessed March 1, 2017.

Khan, U., and Dhar, R. (2006). Licensing Effect in Consumer Choice. *Journal of Marketing Research*, 43, 259-266.

Khan, U., and Dhar, R. (2007). Where There Is a Way, Is There A Will? The Effect of Future Choices on Self-Control. *Journal of Experimental Psychology: General*, *136(2)*, 277-288.

Lee, W. J. *et al.* (2013). You Taste What You See: Do Organic Labels Bias Taste Perceptions? *Food Quality and Preference, 29*, 33-39.

Lindsay, J. (2010) Healthy living guidelines and the disconnect with everyday life. *Critical Public Health, 20(4)* 475-487.

Mankiw, G. (2018). Mankiw's Principles of Economics, 8th edition. Boston, Massachusetts: South-Western College Pub, Chapter 1, pp. 4-6.

McMillan, T. (2015, September 19). The U.S. Doesn't Have Enough Of The Vegetables We're Supposed To Eat. *NPR the Salt*. Retrieved from: http://www.npr.org/sections/thesalt/2015/09/19/441494432/the-u-s-doesnt-have-enough-of-the-vegetables-were-supposed-to-eat. Accessed March 22, 2017.

Nelson, P. (1970). Information and Consumer Behavior. *Journal of Political Economy, 78*(2). 311-329.

Newport, F. (2000, March 6). *Forty Percent of Americans Who Use Glasses Would Consider Laser Eye Surgery*. Retrieved from: http://www.gallup.com/poll/3115/forty-percent-americans-who-use-glasses-would-consider-laser-eye-surgery.aspx. Accessed February 25, 2017.

Office of Disease Prevention and Health Promotion (2015). *Dietary Guidelines for Americans 2015-2020, eighth edition*: Chapter 2: Shifts Needed to Align With Healthy Eating Patterns. Retrieved from: https://health.gov/dietaryguidelines/2015/guidelines/chapter-2/current-eating-patterns-in-the-united-states/. Accessed April 27, 2017.

Posey, K. G. (2016). *Household Income: 2015. American Community Survey Briefs.* Retrieved from U.S. Census website: https://www.census.gov/content/dam/Census/library/publications/2016/demo/acsbr15-02.pdf. Accessed 3 April, 2017.

President's Council on Fitness, Sports, & Nutrition. *Facts and Statistics*. Retrieved from: https://www.fitness.gov/resource-center/facts-and-statistics/. Accessed April 4, 2017.

Roe, B *et al.* (1999). The Impact of Health Claims on Consumer Search and Product Evaluation Outcomes: Results from FDA Experimental Data. *Journal of Public Policy & Marketing 18*(1), 89-105.

Swain, D. P. (Ed)., (2013). *ACMS's Resource Manual for Guidelines for Exercise Testing and Prescription*. Baltimore, MD: Lippincott Williams & Wilkins.

Tal, A. and Wansink, B. (2014). Blinded with science: Trivial graphs and formulas increase ad persuasiveness and belief in product efficacy. *Public Understanding of Science*. 1-9.

Tierney, J. (2011, August 17). Do You Suffer From Decision Fatigue?. *The New York Times Magazine*. Retrieved from: https://www.nytimes.com/section/magazine**.** Accessed January 25, 2017.

United States Census Bureau. (2015, December 22). North Carolina Becomes Ninth State With 10 Million or More People, Census Bureau Reports. *United States Census Bureau*. Retrieved from: https://www.census.gov/newsroom/press-releases/2015/cb15-215.html. Accessed April 3, 2017.

U.S. Census Bureau (2015). *Income in the Past 12 Months (in 2015 Inflation-adjusted Dollars) American Community Survey – 1 Year Estimates.* Retrieved from the American FactFinder on the United States Census Bureau website: https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_15_1YR_S1901&prodType=table. Accessed February 10, 2017.

U.S. Food & Drug Administration. (2016). *Title 21—Food and Drugs Chapter I—Food and Drug Administration Department of Health and Human Services Subchapter B—Food for Human Consumption.* Retrieved from U.S. Food & Drug Administration website: https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?fr=101.65. Accessed March 8, 2017.

U.S. Food & Drug Administration (2016, September 14). *"Natural" on Food Labeling.* Retrieved from: https://www.fda.gov/Food/GuidanceRegulation/GuidanceDocumentsRegulatoryInformation/LabelingNutrition/ucm456090.htm. Accessed April 5, 2017.

Wertenbroch, K., and Dhar, R. (2000). Consumer Choice between Hedonic and Utilitarian Goods. *Journal of Marketing Research, 37*(1), 60-71.