

Obesity and the Consumption Underestimation Bias

PIERRE CHANDON

BRIAN WANSINK*

February 2, 2005

*Pierre Chandon is Visiting Assistant Professor of Marketing at the Kellogg School, Northwestern University, and Assistant Professor of Marketing at INSEAD, Boulevard de Constance, 77300 Fontainebleau, France. Tel: +33 1 60 72 49 87, Fax: +33 1 60 74 61 84, e-mail: pierre.chandon@insead.edu. Brian Wansink is the Julian Simon Research Fellow and Professor of Business Administration, of Nutritional Science, of Advertising, and of Agricultural and Consumer Economics at the University of Illinois, Urbana-Champaign, 350 Wohlers Hall, Champaign IL 61820-6980 USA, e-mail: wansink@uiuc.edu.

ABSTRACT

We contend that people's perception of the number of calories in a meal follows a compressive psychophysical power function and therefore becomes less sensitive as the number of calories of the meal increases. We show, mathematically and empirically, that this psychophysical model solves two enigmatic cross-disciplinary findings: (1) why people systematically underestimate their consumption and (2) why this consumption underestimation bias is more pronounced for overweight people. Five experiments show strong support for the hypothesized model across pre- and post-intake estimations, chosen and nonchosen meals, single and repeated estimations, and separate and joint estimations. We also find that the psychophysical relationship between the estimated and actual number of calories of a meal is independent of the respondent's carelessness, debiasing knowledge, impression management, or estimation anchor—four explanations that have been invoked to explain consumption estimation biases. Our finding that calorie underestimation is caused by meal size and not body size leads to innovative strategies to improve consumption quantity estimations.

According to the National Center for Health Statistics, 65% of U.S. adults are either obese¹ or overweight and the obesity epidemic costs \$78.5 billion annually in adult medical expenditures (Hedley et al. 2004). One factor contributing to increasing obesity rates in both America and abroad is the well-documented tendency that people have to underestimate the number of calories that they consume per meal (Brownell and Battle Horgen 2003). For example, a mere 5% underestimation in a 2000 calorie-a-day diet (e.g., the number of calories contained in a glass of juice) can lead a person to gain up to 10 extra pounds during a year (Newman 2004). In contrast to the often volitional reasons invoked to explain this bias, we develop a psychophysical model of calorie estimation that explains the pattern of the bias and offers innovative implications for improving the accuracy of consumption estimations.

The general consumption estimation bias is robust across self-reported consumption, 24-hour recall interviews, and food frequency questionnaires (Tooze et al. 2004). In a meta-analysis of 43 studies using the doubly labeled water (DLW) technique² to objectively measure actual caloric intakes, Livingstone and Black (2003) find a strong underestimation of consumption (–20% on average), but wide variations across studies (–55% to +40%). This underestimation bias is even more extreme for overweight people (Livingstone and Black 2003; Tooze et al. 2004). Although overweight people eat significantly more than do people with normal weights, they often report eating *less* than do people with normal weights (Shah and Jeffery 1991).

In consumer research, considerable progress has been made in understanding how people process nutrition information (Andrews, Netemeyer, and Burton 1998; Balasubramanian and Cole

¹ Following the guidelines of the World Health Organization, persons are classified as overweight if their body mass index (BMI) is greater than 25 and obese if their BMI is greater than 30. BMI is computed as the ratio of weight, measured in kilos, to squared height, measured in meters.

²The DLW technique involves the administration of water that contains enriched quantities of the “labeled” isotopes deuterium (²H) and oxygen-18 (¹⁸O). The difference in the elimination rate between these two isotopes is measured by urine samples taken on the first and last day of the study (typically, a two-week interval). These quantities yield a measure of carbon dioxide, which can be used to calculate energy expenditure. If weight is maintained during the study period, energy expenditure is equal to energy intake (Trabulsi and Schoeller 2001).

2002; Burton and Creyer 2004; Moorman et al. 2004), estimate single consumption occasions (Krider, Raghbir, and Krishna 2001; Raghbir and Krishna 1999; Wansink and Van Ittersum 2003), and handle health decisions (Moorman 2002; Moorman and Matulich 1993). However, we still lack a good understanding of why people, especially those who are overweight, underestimate the number of calories contained in the meals they consume. This is a critical question for epidemiologists, nutrition researchers, and clinicians: “more fundamentally still, we need to understand why people misreport food intake” (Livingstone and Black 2003, p. 915S), “the reason for the association between BMI and underreporting is unclear” (Tooze et al 2004, p. 803), “our inability to obtain good information on food intake [is] a dilemma for nutrition [but also] an enigma for psychology” Blundell (2000, p. 3).

To address these questions, we develop and test a psychophysical model of calorie estimation that offers a parsimonious, common explanation for the general underestimation bias and for its association with obesity. Drawing on psychophysics research, we hypothesize that subjective caloric content follows a mathematically predictable compressive power function (i.e., a power function with an exponent lower than 1). That is, people’s subjective experience of the number of calories of a meal becomes less sensitive as the number of calories of the meal increases. As a result, people tend to overestimate the number of calories of small meals and fail to notice the increase in the number of calories as the meals become larger, such that the underestimation bias, even when measured as a percentage of the actual number of calories, increases as the size of the meal increases. This suggests that the higher underestimation of overweight consumers might be a spurious consequence of their tendency to consume larger meals. Once the natural confound between body size and meal size is eliminated, our model predicts that overweight and regular persons will have similar consumption estimation biases. In short, calorie underestimation is caused by meal size, not body size

This manuscript is organized as follows: We first develop a psychophysical model of calorie estimation and mathematically show that it can explain the consumption underestimation bias and the higher bias of overweight people. In Experiment 1, we show that people's estimations of the number of calories in a fast-food meal that they just consumed fit with the hypothesized psychophysical model. Experiment 1 also shows that the difference in bias between overweight and regular weight people disappears when the size of meal size is controlled for. In the following four experiments, we replicate these findings in different field and laboratory contexts (pre-/post-intake estimations, chosen/nonchosen meals, single/multiple observations, and separate/joint estimations). We also show that the psychophysical bias is independent of the four possible explanations evoked in the nutrition and epidemiology literature: the respondent's carelessness (Experiment 2), lack of debiasing knowledge (Experiment 3), impression management (Experiment 4), and reliance on outdated estimation anchors (Experiment 5). In the general discussion, we show that understanding the psychophysical basis of consumption estimation biases provides some innovative strategies to help consumers, consumer researchers, nutritionists, and epidemiologists better estimate actual consumption quantity.

A PSYCHOPHYSICAL MODEL OF CALORIE ESTIMATION

To estimate the number of calories consumed during a meal, consumers cannot simply read nutritional labels, which usually are either absent or hard to find (e.g., provided on discarded containers, not displayed in a restaurant). In addition, people often fail to pay attention to caloric content information, even when food labeling is available, but prefer instead to rely on their own estimations (Balasubramanian and Cole 2002). To help us understand how people estimate the caloric content of a meal, we draw from a large body of research in psychophysics, which focuses on the relationship between objective magnitudes and people's subjective experience of them.

Relating Objective Magnitudes and Subjective Experience

The “empirical law of sensation” was established by Stevens (1986) but first documented by Plateau (1872), who discovered that a particular shade of grey always appeared to look “midway” between black and white regardless of whether it was viewed in bright sunshine or in dim light.

Stevens Law, as it has come to be known, can be expressed as follows:

$$(1) \quad \Delta I/I = n\Delta S/S.$$

Equation 1 means that a percentage change in objective magnitude leads to the same percentage change in subjective magnitude (i.e., equal ratios of intensity are associated with equal ratios of sensations). The empirical law of sensation predicts that the subjective impact of adding 100 calories to a meal, for example, depends on the size of the meal, but the subjective impact of doubling the number of calories of a meal is constant, regardless of its size.

Replacing the finite differences in Equation 1 with their corresponding differentials and integrating both sides yields $\ln(S) = n\ln(I) + \ln(k)$, where k is a constant of integration. The resulting psychophysical function (the mathematical relation between objective magnitudes and people’s subjective experience of them) is the following power function:

$$(2) \quad S = aI^n,$$

where S is the subjective magnitude (or sensation), I is the objective magnitude (or intensity), and a is a scaling parameter. The exponent n captures the concavity of the power function (i.e., whether subjective magnitudes are compressive or expansive). This latter definition is easy to perceive; if I is multiplied by r , S is multiplied by r^n . When $n = 1$, subjective magnitudes grow as fast as objective magnitudes. In the special case when $a = n = 1$, estimations are unbiased (people are as likely to overestimate as they are to underestimate the actual magnitude). If $n < 1$, subjective magnitudes increase at a slower rate than do actual magnitudes, and consumers are more likely to underestimate magnitudes as they increase. The reverse occurs if $n > 1$, in which case estimations grow more rapidly than actual magnitudes and consumers become more likely to overestimate

quantities as they increase. From Equation 2, we can derive the intensity level at which sensations are on average accurate ($I^* = a^{1/(1-n)}$), which also represents the magnitude toward which sensations are compressed if $n < 1$.

Calorie Estimations as a Compressive Power Function of Actual Number of Calories

With few exceptions (such as with the perception of the intensity of electric shocks), sensations are always compressive (Poulton 1989). As Krueger (1989) states, “the true psychophysical function is approximately a power function whose exponent normally ranges from 0 to 1, and exceeds 1 only in rather special cases.” For example, people perceive that a second candle adds less brightness than a first candle. Similarly, we also expect that calorie estimations will follow a compressive power function of the actual number of calories. This expectation implies that (1) the range of sensations is smaller than the range of stimuli intensity and (2) sensations increase at a slower rate than intensities. According to the first condition, the accuracy level I^* belongs to the range of observed intensities (i.e., $\min(I) < I^* < \max(I)$), so that large intensities (greater than I^*) are compressed downward toward I^* , whereas small intensities (less than I^*) are compressed upward toward I^* . The second condition implies that the exponent of the power function is smaller than 1 ($n < 1$). Both are testable predictions.

We expect that our hypothesis will hold for calorie estimations obtained in absolute estimation tasks, in which consumers directly estimate the number of calories of a meal that they can see in front of them. . Such a context would minimize a participant’s reliance on their memory and therefore reduces the biases caused by memory-based estimation strategies (Menon and Raghurir 1995). Absolute magnitude estimation tasks also produce fewer context and response biases than do traditional relative magnitude estimation tasks, which ask people to compare a stimuli’s intensity with that of a standard (Gescheider, Bolanowski jr., and Verillo 1992). We use calories instead of other quantity measures because we can then compare estimations across a variety of

foods and beverages that otherwise would need to be measured with different units (e.g., ounces, grams, liters, fluid ounces). In addition, the number of calories in any food is large enough that the biases cannot be driven by left-truncation at 0 (unlike the number of units or servings). Finally, unlike other measures of volume, calorie measurement employs a metric ratio scale that is familiar, easy to use, and directly meaningful for dieting decisions.

We also expect that the hypothesized psychophysical function, because it is driven by low-level perceptual processes, holds in different contexts, as is supported by the well-established robustness of power laws in psychophysics research (Krueger 1989). This robustness is also necessary to explain the magnitude of the biases found in the literature. For example, if pre-intake estimations were refuted by post-intake sensory experiences, consumers should eventually learn to correct their biases. We therefore expect that calorie estimations follow a compressive power function of the actual number of calories whether the estimations are made pre- or post-intake, for chosen or randomly assigned meals, for single or repeated estimations, and while evaluating meals separately or comparing their calorie content.

Implications for Consumption Estimation Biases

Biases in calorie estimation are measured as either the percentage deviation from actual caloric content (*PDEV*) or the log ratio of estimated to actual calories (*LN RATIO*). Both measures are closely related ($LN RATIO = \ln(S/I)$, and $PDEV = (S-I)/I$; hence, $LN RATIO = \ln(PDEV + 1)$). We use the more intuitive measure, *PDEV*, to quantify the magnitude of bias in descriptive analyses but use *LN RATIO* in the modeling analyses because, unlike *PDEV*, it is a linear function of the logarithm of actual consumption quantity if Equation 2 holds (if $S = aI^n$, $LN RATIO = \ln(aI^n/I) = \alpha + \beta \cdot \ln(I)$, where $\alpha = \ln(a)$ and $\beta = (n - 1)$). The derivatives of *PDEV* and *LN RATIO* with respect to *I* are both negative if $n < 1$ ($d(LN RATIO)/d(I) = (n - 1)/I$, and $d(PDEV)/d(I) = (n - 1)I^{n-2}$, which are both negative if $n < 1$). Therefore, both measures of bias decrease (i.e., the

underestimation bias becomes stronger) as actual consumption quantity increases. If, in addition, $\min(I) < I^* < \max(I)$, the estimated number of calories is greater than the actual number of calories for small meals and becomes smaller than the actual number of calories for meals above the accuracy level I^* .

--- Insert Figure 1 about here ---

To illustrate how estimation biases reverse as the number of calories increases, we present Figure 1A, which plots estimated calories S as a function of actual calories I for psychophysical functions with varying degree of compression. As Figure 1A shows, estimated consumption starts above the accuracy line and then drops below it as the number of calories increases. Figure 1A also shows that, as the power exponent n decreases, calorie estimations curves flatten (i.e., estimations are less sensitive to actual consumption quantities), and the number of calories at which estimations are accurate decreases (in Figure 1, $I^* = 1375$ cal when $n = .6$, and $I^* = 124$ cal when $n = .4$). Figure 1B shows the same effects on estimation biases (measured by $LNRATIO$). When calorie estimations follow a compressive power function, $LNRATIO$ is a downward-sloping linear function of the actual number of calories (measured in logs) and the slope is steeper when the degree of compression is higher (i.e., when n is small).

The empirical finding that overweight people underestimate their consumption more than do regular weight people appears directly in Figure 1. Overweight persons tend to consume larger meals than regular weight persons (Subar et al. 2003), which means their estimations pertain to larger meals, which in turn are more likely to be underestimated, because estimations are compressive than are the smaller meals consumed by people with regular weights. However, as Figure 1 shows, the stronger estimation bias of overweight persons can occur even if they rest on the same psychophysical curve as regular weight people. In other words, no intrinsic differences between overweight and regular weight persons are required to explain this empirical finding. This

leads us to the testable prediction that, when the natural confound between people's body mass and consumption quantity is eliminated, overweight and regular weight consumers should exhibit similar consumption estimation biases. For example, Figure 1 shows that if both groups are asked to estimate the same meal—not the distinct meals they normally would choose—their estimates of the number of calories it contains would be similar.

Finally, Figure 1 shows why most studies find a strong consumption underestimation bias, whereas only a minority find consistent overestimation biases. In natural settings, when a broad cross-section of people are surveyed, the distribution of actual consumption quantity is right skewed (most observations are below average but a few large values increase the average actual consumption quantity, Dwyer, Picciano, and Raiten 2003). In these circumstances, the few very large consumption quantities that skew the mean actual consumption quantity are strongly compressed and do not skew the average estimated consumption as much. As a result, the average estimated consumption is lower than the average actual consumption. However, when only a subset of people or consumption occasions with small consumption quantities are surveyed, the reverse finding may appear. For example, Livingstone and Black's (2003) review identifies three groups with a tendency to overestimate or accurately estimate consumption: persons with a very low BMI, children aged 6–12 years, and parents estimating the consumption of their children aged 1–6 years. A common characteristics of these three groups is that they consume smaller quantities than the average person, thus leading to the absence of a general underestimation bias.

EXPERIMENT 1: ESTIMATING CALORIES OF A FAST-FOOD MEAL

The objective of Experiment 1 is to test our central hypothesis that calorie estimations follow a compressive power function of the actual number of calories in the natural setting of fast-food restaurants. Experiment 1 also tests our prediction that the higher biases of overweight consumers are due to their choice of meals that have higher caloric content.

Method

Trained interviewers approached every fourth person as they were finishing their meals in food courts operated by different fast-food restaurants in three medium-sized Midwestern U.S. cities and asked them if they would answer some brief questions for a survey. No mention was made of food at that time. Of the 200 people who were approached, 147 (73.5%) agreed to participate. They were first asked what they ate and drank for their meal and to estimate the number of calories contained in their entire meal. They then answered a few questions about their eating habits and provided their height (in feet and inches) and weight (in pounds), which were used to compute their BMI. During this process, the interviewer unobtrusively recorded and confirmed the type and size of the food and drinks from the packages left on the tray. Nutrition information provided by the fast-food restaurants then was used to compute the actual number of calories of each person's meal.

Unlike DLW, direct observation by trained field researchers allows for unobtrusive measurements of actual consumption quantities in the field but still yields accurate measures of dietary intake (Hise et al. 2002). Of the 147 respondents, 91 were classified as regular weight and 56 as overweight on the basis of their self-reported height and weight.

Results

On average, participants underestimated the caloric content of their meal ($M_S = 613$ cal versus $M_I = 838$ cal, $PDEV = -17.5\%$, $t = -7.45$, $p < .001$). However, this average underestimation hides large differences that depend on the number of calories of the meal. After we dichotomized meals with a median split, we found that the number of calories of small meals was more accurately estimated ($M_{S(\text{small meals})} = 484$ cal versus $M_{I(\text{small meals})} = 508$ cal, $PDEV_{(\text{small meals})} = -.55\%$, $t = -.1$, $p = .92$), whereas the caloric content of large meals was strongly underestimated ($M_{S(\text{large meals})} = 743$ cal versus $M_{I(\text{large meals})} = 1173$ cal, $PDEV_{(\text{large meals})} = -34.59\%$, $t = -10.4$, $p < .001$). Respondents

estimated that larger meals contained, on average, 259 more calories than did smaller meals, though in reality, they contained 665 more calories, more than twice the estimated amount. We illustrate these results in Figure 2A, which shows that estimations tend to be above the accuracy line for small meals but grow more slowly than actual caloric content, so that they quickly fall below the accuracy line as consumption quantities increase.

--- Insert Figure 2 here ---

Figure 2B shows that the average estimation bias (measured by $LNRATIO$) for each quartile of actual quantity forms a straight line, which suggests that consumption estimations follow a power function of actual consumption. To test this possibility directly, we estimated the power model shown in Equation 1. The power model fit the data well ($R^2 = .28$, $F(1, 145) = 55.4$, $p < .01$), and as we show in Figure 2B, the predicted values fall well within the confidence intervals. To examine whether the power model fit the data better than a linear model, we computed the mean average percentage error ($MAPE$) of the power model and of a linear model ($R^2 = .23$, $F(1, 145) = 43.7$, $p < .01$). As we expected, the power model fit significantly better than the linear model ($MAPE_{(Power\ model)} = .38$ versus $MAPE_{(Linear\ model)} = .43$, $t = 4.32$, $p < .01$). The best-fitting power model is $S = 17.0 * I^{.52}$. As expected, the exponent of the model is significantly lower than 1 (t -test of difference from 1 = -6.76 , $p < .001$), and the intersection point is within the range of intensities ($Min(I) = 180\text{ cal} < I^* = 384\text{ cal} < Max(I) = 1921\text{ cal}$). Consistent with our hypothesis, these results show that calorie estimations are compressive, in support of our hypothesis.

Biases for Overweight and Regular Weight Participants. As we expected, the underestimation bias is more pronounced for overweight (OW) than for regular weight (RW) consumers ($PDEV_{OW} = -30.38\%$ versus $PDEV_{RW} = -9.5\%$, $F(1, 145) = 8.90$, $p < .001$). Estimations by overweight consumers were not statistically different from those of regular weight consumers ($M_{S(OW)} = 634$ versus $M_{S(RW)} = 600$, $F(1, 145) = .44$, $p = .51$), though their actual consumption was 266 calories

higher ($M_{I(OW)} = 1003$ versus $M_{I(RW)} = 737$, $F(1, 145) = 17.0$, $p < .001$). What explains this result? Because they eat larger meals, the estimations of overweight consumers (black dots) tend more toward the right in Figure 2A than the estimations of regular weight participants (white dots). They are therefore on the flatter part of the estimation curve, for which estimations are much smaller than is accurate. As we show in Figure 2B, the biases (measured by *LN*RATIO) of overweight and regular weight consumers fall on the same line. This shows that, when the number of calories of a meal is controlled for, overweight and regular weight consumers make the same estimation mistakes.

To test the hypothesis that the differences between overweight and regular weight consumers are mediated by the number of calories in their meals, we estimated two structural equation models. In the first, we test the direct effect of a consumer's overweight status (captured by a binary variable *OW*) on estimation biases (measured by *LN*RATIO). When entered alone, *OW* is a significant predictor of *LN*RATIO ($B = -.26$, $t = -3.2$, $p < .001$). We then estimate the second model, in which *OW* has both a direct effect on *LN*RATIO and an indirect effect through its effect on actual consumption quantity (*I*, measured in logs), which itself influences estimation biases. As we expected, the direct effect of *OW* on *LN*RATIO is not statistically significant ($B = -.12$, $t = -1.6$, $p = .11$), but *OW* has an indirect effect on *LN*RATIO because it increases $\ln(I)$ ($B = .31$, $t = 3.74$, $p < .001$), which itself decreases *LN*RATIO ($B = -.44$, $t = -6.0$, $p < .001$). These results show that, after we control for meal size, there are no differences in the estimation biases of overweight and regular weight participants.

Discussion

Experiment 1 shows a strong consumption underestimation bias in the context of a single meal, minutes after the meal is finished and when consumers know that their estimates can be checked. These results are important because they replicate those of DLW studies and disconfirm

some researchers' expectations that biases will disappear when consumers estimate simple meals or when they believe that their estimations can be readily checked for accuracy (Muhlheim et al. 1998). In addition, Experiment 1 provides strong support for our hypothesis that calorie estimations are a compressive power function of the actual number of calories. Moreover, it shows that biases in consumption estimations are stronger for larger meals, even when measured relative to total quantity, and that the stronger underestimation of overweight consumers is entirely due to the larger meals they consume. Unlike DLW studies, which operate across a longer time period, the setting of Experiment 1 precludes the alternative explanation that the size of the meal explains differences in body mass between overweight and regular weight participants.

The main limitation of Experiment 1 is that it does not rule out carelessness as a driver of the results. The burden of recording and estimating consumption may lead people to omit food occasions or alter their reported intake for convenience, especially when consumption quantity is high (Livingstone and Black 2003). Participants in Experiment 1 might not have been sufficiently motivated to engage in the effortful estimation strategies that would have produced unbiased estimations. In addition, the larger meals may have been more difficult to estimate because they included more food items (such as desserts) or harder-to-estimate items (such as multi-component sandwiches), not just because of their larger number of calories. Finally, because of self-selection, Experiment 1 may have included too many participants who cared little about nutrition or the accuracy of their estimations.

Existing evidence on the effects of carelessness is limited and mixed (Macdiarmid and Blundell 1997; Mela and Aaron 1997). Heymsfield et al. (1995) find that participants' self reports were more accurate for a single meal (for which 86% of actual intake was reported) than for a more burdensome two-week period (for which only 52% of actual intake was reported). Still, these results show that consumption underestimation persists even when estimating and reporting

consumption is easy. It is therefore important to directly test whether respondent's carelessness can explain the results of Experiment 1.

EXPERIMENT 2: EFFECTS OF ESTIMATION CARELESSNESS

The objective of Experiment 2 is to test whether the results of Experiment 1 are caused by respondents' carelessness in remembering or in estimating some of the food they consumed, especially for larger meals. Because psychophysical biases are driven by low-level perceptual processes, we expect that the degree of compression is independent of carelessness. In other words, we expect that carelessness has only a main effect on calorie estimations (and estimation biases) but does not interact with the effects of the actual number of calories.

Method

In Experiment 2, we used the same procedure as in Experiment 1, except that the care with which respondents recorded and estimated the number of calories of their meal was measured. We first asked participants: "Did you pay attention to nutritional information here?" and provided two possible answers (yes or no). We also asked respondents to rate their agreement with two sentences ("I watch what I eat" and "I pay attention to how much I eat") on a nine-point scale anchored at 1 ("strongly disagree") and 9 ("strongly agree"). Averaging their answers to these three questions produced a reasonably reliable scale ($\alpha = .67$). We then categorized respondents as relatively careful or careless through a median split. Because ours is a null hypothesis, we minimized the risk of type II error by collecting data from a greater number of respondents than we did in Experiment 1 ($n = 352$ participants).

The second difference is that we did not ask respondents to report their height or weight. This omission enables us to explore whether the results of Experiment 1 hold in a context that generates less motivation to engage in impression management (we carefully examine the effects of impression management in Experiment 4). To verify this, we asked 153 consumers similar to

those who participated in Experiments 1 and 2 to rate the extent to which someone like them, participating in each study, would be concerned about being evaluated unfavorably and motivated to answer the estimation questions accurately.³ As we expected, Experiment 2 elicited less fear of negative evaluation than did Experiment 1 ($M_{(Experiment\ 2)} = 6.6$ versus $M_{(Experiment\ 1)} = 7.8$, $F(1,151) = 22.2$, $p < .001$), as well as more motivation to answer the estimation questions accurately ($M_{(Experiment\ 2)} = 4.9$ versus $M_{(Experiment\ 1)} = 4.0$, $F(1,151) = 6.0$, $p < .05$). The increase in the response rate (from 73% to 82%, $\chi^2 = 4.5$, $p < .05$) also attests to the lower social demands of Experiment 2.

Results

Replications. The results of Experiment 1 were replicated. On average, participants underestimated the number of calories of their meal ($M_S = 618$ cal versus $M_I = 830$ cal, $PDEV = -18.53\%$, $t = -8.7$, $p < .01$), though as in Experiment 1, the underestimation bias was smaller for small meals ($M_{S(small\ meals)} = 485$ cal versus $M_{I(small\ meals)} = 508$ cal, $PDEV_{(small\ meals)} = -7.18\%$, $t = -2.1$, $p < .05$) than for large meals ($M_{S(large\ meals)} = 781$ cal versus $M_{I(large\ meals)} = 1152$ cal, $PDEV_{(large\ meals)} = -29.88\%$, $t = -12.8$, $p < .001$), and the difference in the magnitude of the bias between small and large meals was statistically significant ($F(1, 350) = 30.8$, $p < .001$). Also as in Experiment 1, the power model fit the data better than a linear model (power model: $R^2 = .34$, $F(1, 350) = 178.7$, $p < .001$; linear model: $R^2 = .31$, $F(1, 350) = 154.2$, $p < .001$), and the difference in predictive validity between the two models was statistically significant ($MAPE_{(power\ model)} = .48$ versus $MAPE_{(linear\ model)} = .42$, $t = 5.7$, $p < .01$). The best-fitting power model is $S = 5.7 * I^{.69}$. As we expected, the exponent is less than 1 ($t = -5.6$, $p < .01$), and the intersection point is within the

³We measured fear of negative evaluation by averaging the answers to the following four items: “This person would worry about what her answers would say about her,” “This person would find these questions embarrassing,” “This person would feel uncomfortable answering these questions,” and “This person would prefer that no one knew her estimations.” The motivation to estimate consumption accurately was measured with two items: “This person would try hard to be accurate” and “This person would be very motivated to answer accurately.” All measures were anchored at 1 = “strongly disagree” and 9 = “strongly agree.” Both measures are reliable ($\alpha = .93$ and $\alpha = .82$, respectively).

range of intensities ($\text{Min}(I) = 140 \text{ cal} < I^* = 274 \text{ cal} < \text{Max}(I) = 2455 \text{ cal}$), in support of our hypothesis that calorie estimations follow a compressive power function.

--- Insert Figure 3 here ---

Effects of carelessness. As we expected, the underestimation bias was, on average, smaller for more careful participants ($M_{S(\text{careful})} = 545 \text{ cal}$ versus $M_{I(\text{careful})} = 686 \text{ cal}$, $PDEV_{(\text{careful})} = -13.93\%$, $t = -4.6$, $p < .01$) than for less careful consumers ($M_{S(\text{careless})} = 689 \text{ cal}$ versus $M_{I(\text{careless})} = 957 \text{ cal}$, $PDEV_{(\text{careless})} = -22.41\%$, $t = -7.3$, $p < .01$), and the difference in the magnitude of the bias between the two groups was statistically significant ($F(1, 350) = 3.9$, $p < .05$). However, these numbers also show that more careful respondents consumed smaller meals ($F(1, 350) = 42.1$, $p < .001$) than less careful consumers. Therefore, the lower bias of these more careful respondents may be driven by their lower consumption, not by intrinsic differences between the groups. We directly test this hypothesis by estimating the following model:

$$(3) \quad \text{LNRATIO} = \alpha + \beta * \ln(I) + \gamma * \text{CARE} + \delta * \ln(I) * \text{CARE},$$

where *CARE* is a mean-centered binary variable that indicates whether consumers were careful, and $\ln(I)$ is the mean-centered log of the actual number of calories.

Instead of estimating Equation 3, we could estimate the power model in Equation 2, in which estimated consumption is the dependent variable, with an interaction between *CARE* and *I*. However, the advantage of Equation 3 is that it can be represented as a well-known linear model and thus shows the interaction effects more easily than does a multiplicative power model. The coefficient for $\ln(I)$ was, as we expected, negative and statistically different from 0 ($\beta = -1 = -.31$, $t = -5.57$, $p < .001$). Neither of the two other coefficients was statistically significant ($\gamma = .03$, $t = .47$, $p = .64$; $\delta = -.03$, $t = -.28$, $p = .78$). After we control for actual consumption quantity, respondents who care about what and how much they eat have the same biases as respondents who

do not care. We also demonstrate this finding in Figure 3; the point estimates for careful and careless consumers are almost perfectly aligned.

Discussion

Experiment 2 shows that the compression of calorie estimations is independent of the amount of care a person devotes to recording and estimating his or her consumption. It also provides another illustration of the misleading results that emerge when we compare biases across groups that differ in their average consumption quantity. To measure whether careful and careless respondents really have different estimation biases, one needs to compare their estimations for meals of similar quantity.

One limitation of Experiment 2 is that it does not rule out that the results may be driven by respondents' lack of information about the bias. Consumers—and researchers—have no intuition that they typically underestimate consumption quantity at such a high degree, especially for larger meals (Livingstone and Black 2003). In support of this explanation, Livingstone and Black (2003) find less underestimation bias among educated people. However, a related stream of research shows that consumers' biases in the estimation of the volume contained in or consumed from a glass (Folkes and Matta 2004; Raghurir and Krishna 1999; Wansink and Van Ittersum 2003), of surfaces (Kriger et al. 2001), or even distances (Raghurir and Krishna 1996) are mostly automatic and unconscious. Raghurir and Krishna (1996) find that making people aware of the direct-distance bias actually increases the prevalence of the bias. It is therefore important to directly test whether lack of information about estimation biases can explain the results of Experiment 2.

EXPERIMENT 3: EFFECTS OF DEBIASING INFORMATION

The objective of Experiment 3 is to test whether the results of Experiments 1 and 2 may be due to respondents' lack of knowledge about the biasing effects of meal size. Following the same reasoning that we used for carelessness, we expect that, debiasing information will only have a

main effect (no interaction) on calorie estimations and estimation biases. In addition, Experiment 3 enables to test whether the results of Experiments 1 and 2 hold for pre-intake estimations. Because they apply to multiple meals, not just the chosen meal, pre-intake estimation biases are more frequent and influence more consumption decisions than do post-intake estimations. Studying pre-intake estimations also reduces potential confounds due to product taste, meal duration, or meal completion. Finally, we conduct Experiment 3 in a more controlled environment than that of Experiments 1 and 2.

Method

In Experiment 3, we use a between-subjects design in which respondents are assigned to either a control condition (replicating the instructions of Experiments 1 and 2) or a debiasing condition (in which they were warned of the consumption underestimation bias). To rule out potential self-selection biases, we conducted Experiment 3 in a laboratory context and obtain estimations from all participants. To rule out differences in the recording burdens of different sized meals, all meals included the same three items (sandwich, fries, and soda), and only the quantity of each item was varied. To reduce the motivation to engage in impression management, estimations were fully private (i.e., participants wrote down their estimations instead of publicly verbalizing them to the interviewer). In addition, we motivated respondents to provide accurate estimations by telling them that the three most accurate respondents would be publicly announced and would be awarded \$50 gift certificates to a local book store.

Respondents were 120 university students who participated in the study to fulfill course requirements (79 participants in the control condition, 41 in the debiasing condition). In one room, participants viewed nine different items of a fast-food meal disposed in a 3×3 matrix. Each row showed one of the three ingredients of the meal: chicken nuggets, fries, or cola purchased at a local fast-food restaurant. Each ingredient was available in three sizes—A, B or C—that formed

the columns of the matrix. Chicken was available in 3 (A), 6 (B), or 12 (C) nuggets; fries were available in 1.45 oz. (A), 2.90 oz. (B), or 5.8 oz. (C) sizes; and the soda, a regular, full-calorie cola, was available in 10 fl. oz. (A), 20 fl. oz. (B), or 40 fl. oz. (C) sizes. The food items were shown on generic paper plates or glasses with no information about their weight or volume.

Participants were instructed to imagine that they were going to order a chicken nugget meal and asked to indicate which size (A, B, or C) of the sandwich, fries, and beverage they would order. Participants in the debiasing condition were told that “when people estimate the number of calories in the food they select, they nearly always underestimate how many calories are in their food. The larger the meal, the more they underestimate. For instance, for a 300 calorie meal, people are fairly accurate, but if it is a 1500 calorie meal, they tend to underestimate by 30%. Knowing this, what is the total number of calories you think are contained in the meal you selected?”. Participants in the control condition were simply asked: “What is the total number of calories you think are contained in the meal you selected?” Finally, all participants indicated their height and weight. In the control condition, 53 participants were classified as regular weight and 26 as overweight (because the sample is small, we do not distinguish between regular weight and overweight participants in the debiasing condition).

In a manipulation check, 155 consumers similar to those who participated in Experiments 1 and 2 indicated that Experiment 3 elicited less fear of negative evaluation than did Experiment 2 (same measures as in the previous pretest, $M_{(Experiment\ 3)} = 3.6$ versus $M_{(Experiment\ 2)} = 6.6$, $F(1,152) = 94.3$, $p < .001$) and a higher motivation to answer the estimation questions accurately ($M_{(Experiment\ 3)} = 6.8$ versus $M_{(Experiment\ 2)} = 4.9$, $F(1,152) = 29.9$, $p < .001$). The procedure thus was successful in increasing the chances that the results of Experiment 3 are not simply due to impression management or carelessness.

Results

Replications. The results from the participants in the control condition replicated all the results of Experiments 1 and 2. On average, there was a significant underestimation in the number of calories ($M_S = 866$ cal versus $M_I = 991$ cal, $PDEV = -7.65\%$, $t = -2.0$, $p < .05$), albeit less strong than in Experiments 1 and 2, probably because of the great effort we undertook to increase care and reduce impression management. The caloric content of small meals was accurately estimated ($PDEV_{(small\ meals)} = 3.08\%$, $t = .70$, $p = .49$), whereas that of large meals was strongly underestimated ($PDEV_{(large\ meals)} = -26.16\%$, $t = -4.5$, $p < .001$), and the difference was statistically significant ($F(1, 77) = 16.0$, $p < .001$), and as in the previous studies, a power model fit the data better than the linear model (power model: $R^2 = .08$, $F(1, 77) = 7.0$, $p < .01$; linear model $R^2 = .07$, $F(1, 77) = 5.7$, $p < .05$), and the difference in predictive validity between the two models was statistically significant ($MAPE_{(power\ model)} = .30$ versus $MAPE_{(linear\ model)} = .33$, $t = 3.0$, $p < .01$). The best-fitting power model is $S = 71.0 * I^{.36}$. Its exponent is significantly lower than 1 ($n = .355$, $t = -4.8$, $p < .001$), and the intersection point is within the range of intensities ($Min(I) = 445$ cal $< I^* = 742$ cal $< Max(I) = 1780$ cal).

As in Experiments 1 and 2, we found greater underestimation by overweight consumers ($M_S = 902$ cal versus $M_I = 1166$ cal, $PDEV = -17.90\%$, $t = -2.5$, $p < .05$) than by regular weight consumers ($M_S = 848$ cal versus $M_I = 905$ cal, $PDEV = -2.63\%$, $t = -.6$, $p = .56$). When it is the only predictor of $LNRATIO$, the coefficient of the binary variable OW is negative and statistically significant ($B = -.21$, $t = -2.2$, $p < .05$), but when we add the actual number of calories ($\ln(I)$) to the regression, the direct effect of OW on $LNRATIO$ becomes insignificant ($B = -.06$, $t = -.6$, $p = .53$), though OW significantly increases $\ln(I)$ ($B = .25$, $t = 3.6$, $p < .001$), which in turn significantly influences the estimation biases ($B = -.61$, $t = -4.2$, $p < .001$). The estimations of overweight and regular weight consumers therefore are similar when we control for the number of

calories of the meal they select. As we show in Figure 4, the biases of overweight and regular weight consumers in the control condition fall on the same line.

--- Insert Figure 4 here ---

Effects of Debiasing Information. Participants in the debiasing condition selected meals with the same number of calories as did participants in the control condition ($M_{I(debias)} = 985$ cal versus $M_{I(control)} = 991$ cal, $F(1, 118) < .1, p = .91$). However, the estimates were significantly higher in the debiasing condition than in the control condition ($M_{S(debias)} = 1097$ cal versus $M_{S(control)} = 866$, cal, $F(1, 118) = 12.4, p < .001$). That is, the debiasing information was successful in eliminating the consumption underestimation bias and even led to a small overestimation bias ($PDEV_{(debias)} = 17.60\%$ versus $PDEV_{(control)} = -7.65\%$, $F(1, 118) = 11.7, p < .001$). If we consider small and large meals separately, however, debiasing information does not eliminate compression. Even in the debiasing condition, there was more underestimation for large meals than for small meals ($PDEV_{(small\ meals)} = 32.66\%$ versus $PDEV_{(large\ meals)} = -1.65\%$, $F(1, 39) = 6.6, p < .05$), which is consistent with our hypothesis. To test this finding directly, we estimated the following model:

$$(4) \quad LNRATIO = \alpha + \beta * \ln(I) + \gamma * DB + \delta * \ln(I) * DB,$$

where DB is a mean-centered binary variable that indicates whether consumers were in the debiasing ($DB = .64$) or control ($DB = -.36$) conditions, and $\ln(I)$ is the mean-centered log of the actual number of calories. Across both conditions, the coefficient for actual caloric content is negative and strongly statistically different from 0 ($\beta = -.63, t = -5.78, p < .001$). The coefficient for the simple effect of the debiasing manipulation is positive ($\gamma = .24, t = 3.49, p < .001$) and thereby indicates less underestimation overall in the debiasing condition than in the control condition. Finally, the coefficient of the interaction is not statistically significant ($\delta = .05, t = .20, p = .84$), so the degree of compression in calorie estimations is the same in both the control and the debiasing conditions. We demonstrate the main effect and lack of interaction for the debiasing

manipulation in Figure 4 with the parallel lines that represent the model predictions for the control and debiased conditions. (Figure 4 does not report results for regular weight and overweight consumers separately because of the limited number of observations.)

Discussion

Experiment 3 shows the robustness of the findings of Experiments 1 and 2, even when consumers make their estimations before intake and in a context with a 100% response rate, a low and controlled estimation burden, strong incentives for accuracy, and little incentive for impression management. In addition, Experiment 3 shows that providing information about consumption underestimation biases only shifts the mean calorie estimations upward. Despite learning that estimation biases are stronger for larger meals, participants in Experiment 3 continued to underestimate large meals more than they did small meals, and the degree of compression of calorie estimation remained constant.

One limitation of Experiment 3, however, is that it does not eliminate the possibility that the results may be caused by impression management (Leary and Kowalski 1990). Impression management may explain the consumption underestimation bias because of the widespread belief that maintaining a low body weight and eating less are socially rewarded (Muhlheim et al. 1998). It also may explain the greater underestimation by overweight consumers, for whom the perceived social pressure to minimize consumption is highest (Crandall 1994).

Explanations involving impression management have received generally mixed support. Muhlheim et al. (1998) attempt to reduce the motivation to engage in impression management by telling respondents that their two-week self-reports would be verified by DLW. However, this “bogus pipeline” procedure increased self-reported consumption from 55% to only 61% of true quantity. McKenzie et al. (2002) tried to reduce the motivation for impression management by using overweight interviewers, but this method failed to improve estimation accuracy. However,

other studies found no relationship between social desirability and consumption estimation biases (Horner et al. 2002; Tooze et al. 2004). It is therefore important to directly test whether impression management can explain the results of Experiment 3.

EXPERIMENT 4: EFFECTS OF IMPRESSION MANAGEMENT

The objective of Experiment 4 is to test whether the results of Experiments 1–3 might be due to impression management. Following the same reasoning that we provided for the other alternative explanations, we expect that impression management has only a main effect on calorie estimations (and estimation biases) but does not interact with the effects of the actual number of calories. In addition, Experiment 4 enables us to examine whether our results hold when participants make repeated estimations of different meals rather than a single estimation of a chosen meal (as in Experiments 1–3).

Method

In Experiment 4, we recruited 55 students from a university campus and asked them to estimate the number of calories of the same eight meals. This procedure makes meal size independent of people's BMI. If, as we argue, the stronger underestimation biases of overweight consumers are due to their selection of larger meals, there should be no differences between the estimations of overweight and regular weight consumers in Experiment 4, because everyone is estimating the same meals. In addition, we measured the motivation to engage in impression management by measuring social desirability, a personality trait linked to the motivation to engage in impression management (Crowne 1991; Leary and Kowalski 1990).

Each meal consisted of a sandwich, fries, and colas in different quantities, and total calories ranged between 190 and 1480. The eight meals were displayed on separate tables, and the order in which they were viewed was random across participants. Participants provided their height and weight, which enabled us to split them into regular weight ($n = 39$) and overweight ($n = 16$)

groups. They then responded to the 10-item M-C social desirability scale, a shorter but reliable version of the original 33-item M-C social desirability scale, which captures people's need to obtain approval by responding in culturally appropriate and acceptable manners (Barger 2002). The 10-item M-C social desirability scale asks participants to indicate whether each of 10 statements (e.g., "I never hesitate to go out of my way to help someone in trouble") is true or false as it pertains to them personally. Answers are coded 1 if they are socially desirable and 0 if not and then summated. Because we obtained some ties in the M-C scores, we assigned 22 participants to the high social desirability (SD) group and 33 to the low SD group.

In a manipulation check, 160 consumers similar to those who participated in Experiments 3 and 4 indicated that Experiment 4 elicited less fear of negative evaluation than did Experiment 3 (same measures as in the previous pretest, $M_{(experiment\ 4)} = 4.3$ versus $M_{(experiment\ 2)} = 6.6$, $F(1,157) = 63.4$, $p < .001$) and significantly more motivation to answer the estimation questions accurately ($M_{(experiment\ 4)} = 6.5$ versus $M_{(experiment\ 2)} = 4.9$, $F(1,157) = 28.8$, $p < .001$). The social desirability scores of overweight and regular weight participants were almost identical ($SD_{(OW)} = 5.00$ versus $SD_{(RW)} = 5.17$, $F(1, 53) < .1$, $p = .76$), which already suggests that social desirability cannot explain the potential differences between the groups.

Results

Replications. All the results from Experiments 1–3 that pertain to the underestimation bias and the stronger underestimation for larger meals were replicated. There is a significant underestimation of consumption quantity ($M_S = 561$ cal versus $M_I = 700$ cal, $PDEV = -11.4\%$, $t = -4.8$, $p < .001$), but the caloric content of small meals was estimated accurately ($PDEV_{(small\ meals)} = 2.0\%$, $t = .5$, $p = .60$), the caloric content of large meals was strongly underestimated ($PDEV_{(large\ meals)} = -24.8\%$, $t = -10.0$, $p < .001$), and the two values were strongly statistically different ($F(1, 436) = 34.8$, $p < .001$). As in our previous experiments, the power model fit the data better than a

linear model (power model: $R^2 = .65$, $F(1, 436) = 309.9$, $p < .01$; linear model $R^2 = .62$, $F(1, 436) = 273.5$, $p < .001$), and the difference in predictive validity between the two models is statistically significant ($MAPE_{(power\ model)} = .50$ versus $MAPE_{(linear\ model)} = .59$, $t = 8.5$, $p < .001$). The best-fitting power model is $S = 4.0 * I^{.74}$, the exponent is less than 1 ($t = -6.2$, $p < .001$), and the intersection point is within the range of intensities ($Min(I) = 190\text{ cal} < I^* = 205\text{ cal} < Max(I) = 1480\text{ cal}$).

Effects of Body Mass and Social Desirability. As we expected, the average estimation biases of overweight and regular weight respondents are identical ($PDEV_{(RW)} = -11.1\%$ versus $PDEV_{(OW)} = -12.0\%$, $F(1, 436) < .1$, $p = .86$), and as we show in Figure 5A, the estimations of both groups are almost perfectly aligned. If we decouple body mass from consumption quantity, the differences between overweight and regular weight consumers disappear. Also as we expected, the consumption underestimation bias is stronger for respondents with a high social desirability score ($PDEV_{(low\ SD)} = -7.8\%$ versus $PDEV_{(high\ SD)} = -16.7\%$, $F(1, 436) = 3.4$, $p < .07$). However, Figure 5B shows that the predictions for high and low social desirability respondents fall on parallel lines, which suggests that social desirability does not interact with actual consumption quantity. To test this finding directly, we estimated the following model:

$$(5) \quad LNRATIO = \alpha + \beta * \ln(I) + \gamma * SD + \delta * OW + \varepsilon * \ln(I) * OW + \theta * \ln(I) * SD + \lambda * OW * SD,$$

where $\ln(I)$ is the mean-centered logarithm of actual consumption quantity, SD is a mean-centered binary variable that indicates whether consumers have a high score on the M-C social desirability scale ($SD = .60$) or not ($SD = -.40$), and OW is a mean-centered binary variable that indicates whether consumers are overweight ($OW = .71$) or not ($OW = -.29$). Equation 5 incorporates overweight status as a control variable because it was not independently manipulated.

--- Insert Figure 5 here ---

As we expected, the coefficients of actual consumption quantity and social desirability are negative ($\beta = -.26$, $t = -6.26$, $p < .01$; $\gamma = -.15$, $t = -2.93$, $p < .01$). This indicates two main effects:

The underestimation bias is more pronounced for larger meals and with high social desirability biases. Consistent with our expectations, none of the other coefficients was statistically significant ($\delta = .05, t = .81, p = .42$; $\varepsilon = -.06, t = -.62, p = .548$; $\theta = .12, t = 1.4, p = .17$; $\lambda = -.15, t = -1.34, p = .18$). The degree of compression of the actual consumption quantity is independent of social desirability and body mass, and there is no interaction between social desirability and body mass. In addition, these results cannot be explained by a lack of statistical power. Our sample size ($n = 438$) is significantly larger than needed to detect the average effect sizes found in the literature at a .05 (two-tailed) significance level with the conventional .80 power. These sizes are $n^* = 193$ for the association between body mass and estimation biases (given that $r = -.20$, as found by Livingstone and Black 2003) and $n^* = 58$ for the association between social desirability and estimation biases (given $r = -.36$, as found by Taren et al. 1999).

Discussion

Experiment 4 shows that social desirability, a personality trait linked to the motivation to engage in impression management, has only a main effect in reducing the reported number of calories. Social desirability does not change the degree of compression of actual consumption quantity. Experiment 4 also shows that, when all respondents evaluate multiple meals and when body mass is independent of actual consumption quantity, the estimation biases of overweight persons are indistinguishable from those of regular weight consumers. Finally, Experiment 4 shows that consumption estimations are a compressive function of actual consumption quantity, even for nonchosen meals and repeated estimations.

One limitation of Experiment 4 is that it does not rule out the final alternative explanation of our results, namely, that consumers anchor their estimates on the basis of nutritional information that is outdated and therefore too low. For example, Heymsfield et al. (1995) cite a textbook that attributes 200 calories per bagel. In reality, the National Institute of Health estimates that modern

bagels typically have 350 calories and sometimes contain up to 552 calories (Burros 1994). Using nationally representative data, Nielsen and Popkin (2003) find that, between 1977 and 1996, food portion sizes have increased for all categories except pizzas. In addition, they find that the increase is greater for high-calorie foods such as cheeseburgers (from 381 cal to 485 cal) than for lower calorie food such as soft drinks (from 125 cal to 155 cal). This shift may explain the higher underestimation biases of overweight consumers, who are more likely to eat high-calorie food items. It is therefore important to directly test whether estimation anchors can explain the results of Experiment 4.

EXPERIMENT 5: EFFECTS OF ESTIMATION ANCHORS

The objective of Experiment 5 is to test whether the results of Experiments 1–4 are due to estimation anchors. Following the same reasoning as that for the other alternative explanations, we expect that manipulating the numbers that consumers use as self-generated anchors when estimating calories has only a main effect on estimation biases but does not interact with the effects of the actual number of calories. In addition, Experiment 5 enables us to test whether calorie estimations follow a compressive power function when people make joint estimations (comparing two meals) as opposed to separate evaluations of single meals.

We might expect lower biases for joint and relative estimations than for the separate and absolute estimations examined in Experiments 1–4 because they are more frequent (i.e., consumption decisions more often involve comparisons between different meals than between a meal or nothing). Another reason to expect smaller biases is that the difference in the number of calories in different meals typically is smaller than the number of calories of each meal.

Furthermore, prior research has shown that consumers pay more attention to hard-to-evaluate attributes in joint evaluations than in separate evaluations (Hsee 1996), and therefore, people

should be less likely to underestimate the number of calories (a hard-to-estimate attribute) when comparing two meals than when evaluating each meal separately.

Method

In Experiment 5, we use a 2×4 design with two between-subjects conditions (high versus low estimation anchors) and four within-subject conditions (four different meal sizes). We recruited 87 university students on campus, showed them four pairs of meals, and asked them to estimate the difference in the number of calories of the two meals of each pair. In each pair, one meal was a high-calorie meal and the other low calorie. To improve the external validity of the study, we did not follow the typical anchoring procedure, which consists of providing external anchors by asking respondents if the number to be estimated is above or below a certain level. Instead, we followed the procedure used by Epley and Gilovich (2001) to manipulate self-generated estimation anchors. Participants in the high (low) estimation anchor condition were asked to imagine that they had decided to eat the high (low) calorie meal of the pair.

The four pairs of meals were arranged in increasing order of size and identified as A through D. The high-calorie meal from pair A consisted of chicken nuggets, fries, and a regular cola for a total of 370 calories. The low-calorie meal from pair A consisted of half a low-calorie turkey sandwich, baked chips, and diet cola for a total of 190 calories. The number of calories that participants had to estimate was therefore 180. Pairs B, C, and D contained the same meals as pair A, but the size of each meal was increased by a factor of 2, 3, and 4, respectively, compared with the meal in pair A. The number of calorie to be estimated was therefore 360 calories for pair B, 540 calories for pair C, and 720 calories for pair D.

Results and Discussion

Replications. The results of Experiments 1–4 were replicated. Respondents significantly underestimated the number of calories separating the meals of each pair ($M_S = 354$ cal versus M_I

= 450 cal, $PDEV = -9.7\%$, $t = -2.5$, $p < .05$). In addition, the estimation was more accurate when the number of calories separating the two meals was small (for pairs A and B, $PDEV = 11.23\%$, $t = 1.7$, $p = .09$) but was strongly below the actual difference when the difference between the two meals was high (for pairs C and D, $PDEV = -30.6\%$, $t = -8.1$, $p < .001$). The difference in the magnitude of the bias between large and small quantities was strongly statistically different ($F(1, 343) = 18.4$, $p < .001$). As in previous experiments, a power model fit the data better than a linear model (power model: $R^2 = .33$, $F(1, 343) = 42.3$, $p < .01$; linear model $R^2 = .31$, $F(1, 343) = 37.6$, $p < .01$), and the difference in the predictive validity between the two models was statistically significant ($MAPE_{(power\ model)} = .74$ versus $MAPE_{(linear\ model)} = .91$, $t = 10.4$, $p < .001$). The best-fitting power model is $S = 14.6 * I^{.49}$, the exponent is significantly lower than 1 ($t = -6.9$, $p < .001$), and the intersection point is within the range of intensities ($Min(I) = 180\text{ cal} < I^* = 192\text{ cal} < Max(I) = 720\text{ cal}$).

Anchoring Effects. Estimations were accurate in the high anchor condition ($PDEV_{(high\ anchor)} = 6.8\%$, $t = 1.2$, $p = .22$) and significantly below the actual number in the low anchor condition ($PDEV_{(low\ anchor)} = -26.0\%$, $t = -4.9$, $p < .01$). The difference between the two biases was strongly statistically significant ($F(1,344) = 18.4$, $p < .001$). However, the point estimates and model predictions, as shown in Figure 6, are almost perfectly parallel across conditions. Therefore, as we expected, the manipulation of estimation anchors has only a main effect on the estimations. To provide a direct test of this finding, we estimated the following model:

$$(6) \quad LNRATIO = \alpha + \beta * \ln(I) + \gamma * ANCH + \delta * \ln(I) * ANCH,$$

where $ANCH$ is a binary variable that indicates whether consumers were in the high ($ANCH = .5$) or low ($ANCH = -.5$) self-generated estimation anchor condition, and $\ln(I)$ is the mean-centered logarithm of the actual number of calories. As we found in all the previous experiments, the coefficient for $\ln(I)$ was strongly negative ($\beta = -.51$, $t = -7.29$, $p < .001$). Although the simple

effect of *ANCH* was strongly positive ($\gamma = .47, t = 6.05, p < .001$), the interaction effect was not statistically significant ($\delta = .10, t = .73, p = .47$).

Overall, Experiment 5 shows that consumption estimations are a compressive function of actual consumption quantity, even when consumers estimate the difference in the number of calories between two meals. Experiment 5 also shows that manipulating estimation anchors has only a main effect on consumption estimations. When participants were told that they had chosen the high-calorie meal of the pair (high estimation anchor condition), their estimations were almost 50% higher than when they were told that they had chosen the low-calorie meal. However, the degree of compression of these estimations was the same across both conditions.

GENERAL DISCUSSION

Biases in consumption estimations are a significant contributing factor to the obesity epidemic. Considerable evidence in the epidemiology and nutrition literature shows that people underestimate the number of calories they consume and that the consumption underestimation bias is more extreme for overweight people. Yet, despite the importance of understanding the origin of these biases, none of the explanations developed has proved compelling. In this research, we develop a psychophysical model of calorie estimation that explains these enigmatic findings. We argue that people's perception of the number of calories in a meal follows a compressive power function of the actual number of calories (i.e., a power function with diminishing sensitivity). We find strong support for the hypothesized model in five experiments that involve pre- as well as post-intake estimations, chosen as well as nonchosen meals, single as well as repeated observations, and separate as well as joint estimations.

The direction and size of the consumption estimation bias depends on the number of calories of the meal being estimated, even when the bias is measured as a percentage of the true number of calories. The underestimation bias is stronger for increasingly larger meals, and it decreases as the

number of calories to be estimated decreases, to the point that small meals are often overestimated. We also show, mathematically and empirically, that this explains why the consumption underestimation bias is more pronounced for overweight consumers. It is not that overweight consumers are intrinsically worse predictors than regular weight consumers. The higher underestimation of overweight consumers is a spurious consequence of their tendency to consume larger meals. Calorie underreporting is a problem with meal size, not body size.

We also find that the psychophysical relationship between objective and subjective consumption quantity is neither mediated nor moderated by any of the four existing explanations for consumption estimation biases. Respondents' carelessness, a lack of information about the bias, impression management, or anchoring based on outdated nutritional information does not modify or suppress the effects of the objective number of calories on estimation biases. That these four factors all have a main effect on consumption estimations implies that the absence of an interaction between them and the effect of meal size is not due to a lack of power or poor calibration. Rather, our findings show that psychophysical compression is an independent source of consumption estimation biases and therefore supplements existing explanations.

Research Implications

Biases in consumption estimations mean that studies that use uncorrected self-reported consumption data as an independent or dependent variable are biased (Livingstone and Black 2003). Unfortunately, the prohibitive cost of the DLW technique probably means that most researchers in marketing, nutrition, and epidemiology will still need to depend on self-reported data in many instances. How can self-reported consumption data be debiased? One existing technique is to eliminate data from known underreporters (e.g., obese respondents). Unfortunately, this technique eliminates data from those persons who are of the greatest interest. Another technique consists of applying a correction factor to all observations or to different groups of

respondents (such as overweight versus regular weight consumers). Our study suggests that better results can be achieved by applying a different correction factor for each meal size. Consider Experiment 1. Because consumption was underestimated by an average of 17.5%, a general correction factor would be to multiply self-reports by 1.21 ($1/(1 - .175)$). When we compare the accuracy of self-reported data in Experiment 1 corrected using (1) a single correction factor, (2) a different factor for overweight and regular weight consumers, and (3) a different factor for small and large meals, we find that the single-factor correction predicts true consumption as well as the “obesity” correction ($MAPE_{(single\ factor)} = .40$ versus $MAPE_{(overweight\ vs.\ regular\ weight)} = .38$, $t = 1.6$, $p = .10$) but both are outperformed by the meal size correction ($MAPE_{(small\ vs.\ large\ meals)} = .36$, $t = 2.2$, $p < .05$).

Another fruitful area for additional research would be to examine whether consumption estimations can be improved by using different measurement units (such as volume or number of servings) or by focusing people more on external cues (such as the size of the plate or type of food) than on biased internal sensations. Another area worthy of research would be to examine the effects of familiarity and expectations. Our results rule out the idea that consumption estimation biases are due to a lack of familiarity with large meals (overweight people are no better predictors than people with a regular weight who are less used to consuming larger meals). Expectations of a lower-calorie meal might even aggravate the consumption underestimation bias, so that consumers would be more accurate when estimating a typical high-calorie fast-food meal than when estimating an objectively healthier meal. Such an analysis would need to control for the differences in the size of the high- and low-calorie meals, which will create its own biases. Overall, our study shows that any analyses of factors that affect the biases in consumption estimation must carefully disentangle improvements that are due to changes in the quantity consumed from those that are due to intrinsic improvements in estimation accuracy.

Public Policy Implications

One important area for public policy research would be to study the link between biases in calorie estimation and consumption behavior. For example, it would be useful to examine whether people who tend to underestimate the number of calories of a fast-food meal are more sensitive to “supersize” promotions that offer larger meals for a modest price increase than people whose estimations are unbiased. This raises the question of what medical practitioners, clinicians, and health policy professionals can do to improve consumers’ estimations of the number of calories in a meal. Our results show that simply informing people about the bias is ineffective. One solution would be to make nutritional information more salient in restaurants. Whereas, it may not influence estimates, it may influence choice (Burton and Creyer 2004).

A more general solution might be to tell consumers to separate their meals mentally into smaller portions and estimate the number of calories of each portion independently. For example, our model predicts that estimation accuracy would greatly improve if consumers estimated the number of calories of half of their meal and then doubled that number. In Experiment 1, for example, the average meal contained 838 calories, but the predicted estimation was 578 calories ($17 \cdot 838^{.524}$). In contrast, the predicted estimation for a “half meal” of 419 calories was 402 calories ($17 \cdot 419^{.524}$), which, when doubled, is very close to the true number of calories.

REFERENCES

- Andrews, J. Craig, Richard G. Netemeyer, and Scot Burton (1998), "Consumer Generalization of Nutrient Content Claims in Advertising," *Journal of Marketing*, 62 (4), 62-75.
- Balasubramanian, Siva K. and Catherine Cole (2002), "Consumers' Search and Use of Nutrition Information: The Challenge and Promise of the Nutrition Labeling and Education Act," *Journal of Marketing*, 66 (3), 112.
- Barger, Steven D. (2002), "The Marlowe-Crowne Affair: Short Forms, Psychometric Structure, and Social Desirability," *Journal of Personality Assessment*, 79 (2), 286-305.
- Blundell, John E. (2000), "What Foods Do People Habitually Eat? A Dilemma for Nutrition, an Enigma for Psychology," *American Journal of Clinical Nutrition*, 71 (1), 3-5.
- Brownell, Kelly D. and Katherine Battle Horgen (2003), *Food Fight : The Inside Story of the Food Industry, America's Obesity Crisis, and What We Can Do About It*, New York: McGraw-Hill.
- Burros, Marian (1994), "Eating well: Some Bagels are Hefty in Calories," *The New York Times*, July 6.
- Burton, Scot and Elizabeth H. Creyer (2004), "What Consumers Don't Know Can Hurt Them: Consumer Evaluations and Disease Risk Perceptions of Restaurant Menu Items," *Journal of Consumer Affairs*, 38 (1), 121-145.
- Crandall, Christian S. (1994), "Prejudice Against Fat People: Ideology and Self-Interest," *Journal of Personality & Social Psychology*, 66 (5), 882-894.
- Crowne, Douglas P. (1991), "From Response Style to Motive - A Citation-Classic Commentary on The Approval Motive - Studies in Evaluative Dependence by Crowne, D.P., and Marlowe, D.," *Current contents*, 16 (August 5), 18.
- Dwyer, Johanna, Mary Frances Picciano, and Daniel J. Raiten (2003), "Estimation of Usual Intakes: What We Eat in America-NHANES," *Journal of Nutrition*, 133 (2), 609S-623.
- Epley, Nicholas and Thomas Gilovich (2001), "Putting Adjustment Back in the Anchoring and Adjustment Heuristic," *Psychological Science*, 12, 391-396.

- Folkes, Valerie S. and Shashi Matta (2004), "The Effect of Package Shape on Consumers' Judgments of Product Volume: Attention as a Mental Contaminant," *Journal of Consumer Research*, 31 (2), 390-391.
- Gescheider, George A., Stanley J. Bolanowski jr., and Ronald T. Verillo (1992), "Sensory, Cognitive and Response Factors in the Judgment of Sensory Magnitude," in *Psychophysical Approaches to Cognition*, ed. Daniel Algom, Amsterdam, NL: Elsevier, 575-621.
- Hedley, Allison A., Cynthia L. Ogden, Clifford L. Johnson, Margaret D. Carroll, Lester R. Curtin, and Katherine M. Flegal (2004), "Prevalence of Overweight and Obesity Among US Children, Adolescents, and Adults, 1999-2002," *The Journal of the American Medical Association*, 291 (23), 2847-2850.
- Heymsfield, Steven B., Pamella C. Darby, Lauren S. Muhlheim, and Dympna Gallagher (1995), "The calorie: Myth, Measurement, and Reality," *American Journal of Clinical Nutrition*, 62 (5), 1034S-1041S.
- Hise, Mary E., Debra K. Sullivan, Dennis J. Jacobsen, Susan L. Johnson, and Joseph E. Donnelly (2002), "Validation of energy intake measurements determined from observer-recorded food records and recall methods compared with the doubly labeled water method in overweight and obese individuals," *American Journal of Clinical Nutrition*, 75 (2), 263-267.
- Horner, Neilann K., Ruth E. Patterson, Marian L. Neuhouser, Johanna W. Lampe, Shirley A. Beresford, and Ross L. Prentice (2002), "Participant characteristics associated with errors in self-reported energy intake from the Women's Health Initiative food-frequency questionnaire," *American Journal of Clinical Nutrition*, 76 (4), 766-773.
- Hsee, Christopher K. (1996), "The Evaluability Hypothesis: An Explanation for Preference Reversals Between Joint and Separate Evaluations of Alternatives," *Organizational Behavior & Human Decision Processes*, 67 (3), 247-257.
- Krider, Robert E., Priya Raghubir, and Aradhna Krishna (2001), "Pizzas: Pi or Square? Psychophysical Biases in Area Comparisons," *Marketing Science*, 20 (4), 405.
- Krueger, Lester E. (1989), "Reconciling Fechner and Stevens: Toward a Unified Psychophysical Law," *Behavioral and Brain Sciences*, 12, 251-320.

- Leary, Mark R. and Robin M. Kowalski (1990), "Impression Management: A Literature Review and Two-Component Model," *Psychological Bulletin*, 107 (1), 34-47.
- Livingstone, M. Barbara E. and Alison E. Black (2003), "Markers of the Validity of Reported Energy Intake," *Journal of Nutrition*, 133 (3), 895S-920.
- Macdiarmid, J.I. and J.E Blundell (1997), "Dietary under-reporting: What people say about recording their food intake," *European Journal of Clinical Nutrition*, 51 (3), 199-200.
- McKenzie, Debra C., Rachel K. Johnson, Jean Harvey-Berino, and Beth Casey Gold (2002), "Impact of Interviewer's Body Mass Index on Underreporting Energy Intake in Overweight and Obese Women," *Obesity Research*, 10 (6), 471-477.
- Mela, David J. and Jacqueline I. Aaron (1997), "Honest but Invalid What Subjects Say about Recording their Food Intake," *Journal of the American Dietetic Association*, 97 (7), 791-793.
- Menon, Geeta and Priya Raghubir (1995), "Behavioral Frequency Judgments: An Accessibility-Diagnosticity Framework," *Journal of Consumer Research*, 22 (2), 212-228.
- Moorman, Christine (2002), "Consumer Health under the Scope," *Journal of Consumer Research*, 29 (1), 152-158.
- Moorman, Christine, Kristin Diehl, David Brinberg, and Blair Kidwell (2004), "Subjective Knowledge, Search Locations, and Consumer Choice," *Journal of Consumer Research*, 31 (3), 673-680.
- Moorman, Christine and Erika Matulich (1993), "A model of consumers' preventive health behaviors: The role of health motivation and health ability," *Journal of Consumer Research*, 20 (2), 208-228.
- Muhlheim, Lauren S., David B. Allison, Stanley Heshka, and Steven B. Heymsfield (1998), "Do Unsuccessful Dieters Intentionally Underreport Food Intake?," *International Journal of Eating Disorders*, 24 (3), 259-266.
- Newman, Cathy (2004), "Why are we so fat?," *National Geographic*, 206 (2), 46-62.
- Nielsen, Samara Joy and Barry M. Popkin (2003), "Patterns and trends in food portion sizes, 1977-1998" *The Journal of the American Medical Association*, 289, 450-453.

- Plateau, Joseph Antoine (1872), "Sur la Mesure des Sensations Physiques, et sur la Loi qui Lie l'Intensité de ces Sensations à l'Intensité de la Cause Excitante," *Bulletins de l'Académie Royale des Sciences, des Lettres, et des Beaux-Arts de Belgique*, 33 (376-388).
- Poulton, E. C. (1989), *Bias in Quantifying Judgments*, London: Lawrence Erlbaum Associates.
- Raghubir, Priya and Aradhna Krishna (1996), "As the Crow Flies: Bias in Consumers' Map-Based Distance Judgments," *Journal of Consumer Research*, 23 (1), 26-39.
- _____ (1999), "Vital dimensions in volume perception: Can the eye fool the stomach?," *Journal of Marketing Research*, 36 (3), 313-326.
- Shah, Meenah and Robert W. Jeffery (1991), "Is Obesity Due to Overeating and Inactivity, or to a Defective Metabolic Rate? A Review," *Annals of Behavioral Medicine*, 13, 73-81.
- Stevens, Stanley Smith (1986), *Psychophysics: Introduction to its Perceptual, Neural, and Social Prospects*, Oxford, UK: Transaction Books.
- Subar, Amy F., Victor Kipnis, Richard P. Troiano, Douglas Midthune, Dale A. Schoeller, Sheila Bingham, Carolyn O. Sharbaugh, Jillian Trabulsi, Shirley Runswick, Rachel Ballard-Barbash, Joel Sunshine, and Arthur Schatzkin (2003), "Using Intake Biomarkers to Evaluate the Extent of Dietary Misreporting in a Large Sample of Adults: The OPEN Study," *American Journal of Epidemiology*, 158 (1), 1-13.
- Taren, Douglas .L., M. Tobar, A. Hill, W. Howell, C. Shisslak, I. Bell, and C. Ritenbaugh (1999), "The association of energy intake bias with psychological scores of women," *European Journal of Clinical Nutrition*, 53 (7), 570-578.
- Tooze, Janet A, Amy F Subar, Frances E Thompson, Richard Troiano, Arthur Schatzkin, and Victor Kipnis (2004), "Psychosocial predictors of energy underreporting in a large doubly labeled water study," *The American Journal of Clinical Nutrition*, 79 (5), 795-804.
- Trabulsi, Jillian and Dale A. Schoeller (2001), "Evaluation of dietary assessment instruments against doubly labeled water, a biomarker of habitual energy intake," *American Journal of Physiology - Endocrinology and Metabolism*, 281 (5), E891-899.
- Wansink, Brian and Koert Van Ittersum (2003), "Bottoms Up! The Influence of Elongation on Pouring and Consumption Volume," *Journal of Consumer Research*, 30 (3), 455-463.

FIGURE 1

EFFECTS OF THE DEGREE OF COMPRESSION OF THE PSYCHOPHYSICAL FUNCTION ON ESTIMATIONS (PANEL A) AND ESTIMATION BIASES (PANEL B)

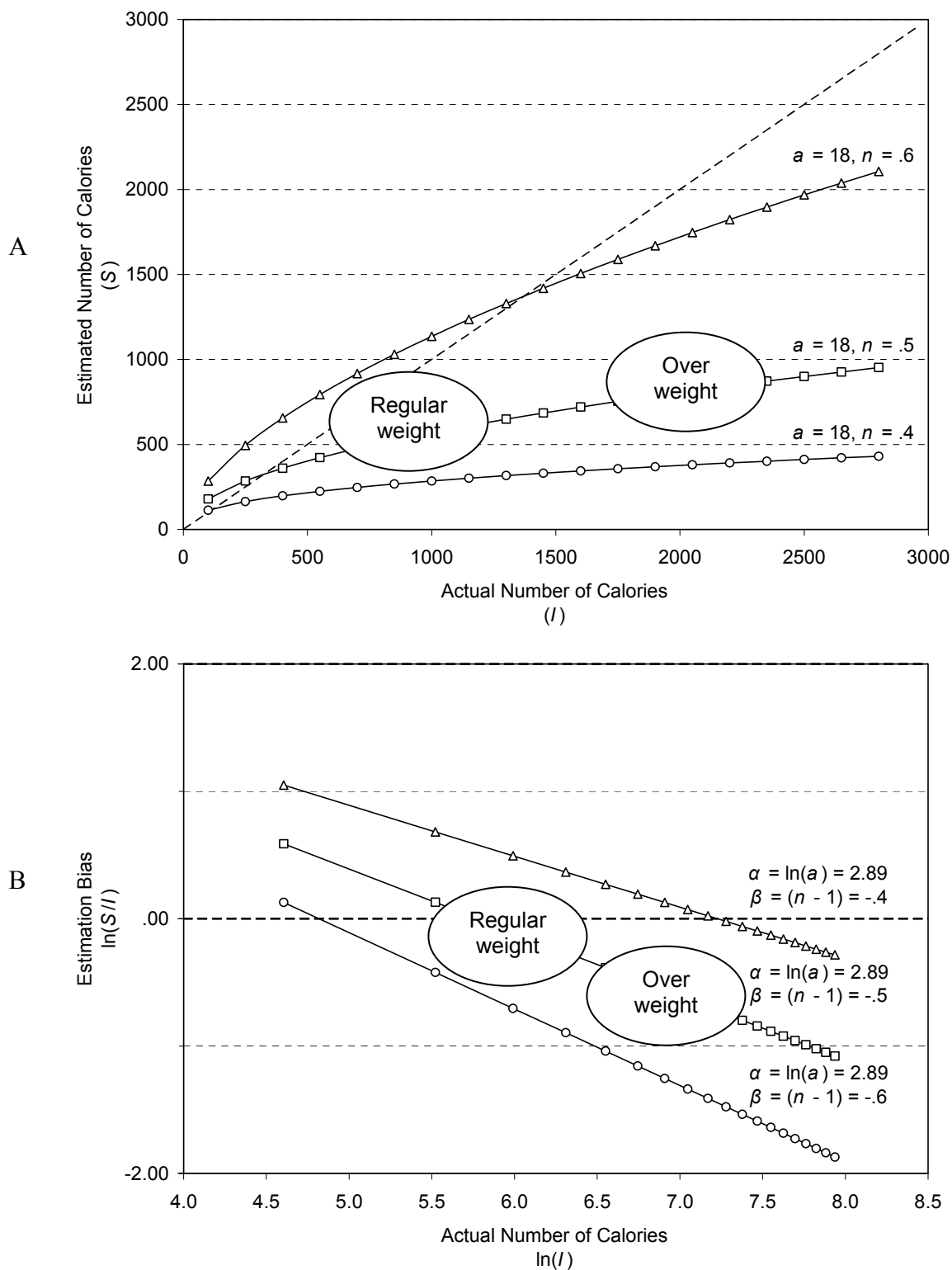


FIGURE 2

EXPERIMENT 1: ESTIMATED VERSUS ACTUAL NUMBER OF CALORIES (PANEL A)
AND PREDICTED AND OBSERVED (MEAN AND 95% CI) ESTIMATION BIASES AS A
FUNCTION OF ACTUAL NUMBER OF CALORIES (PANEL B)

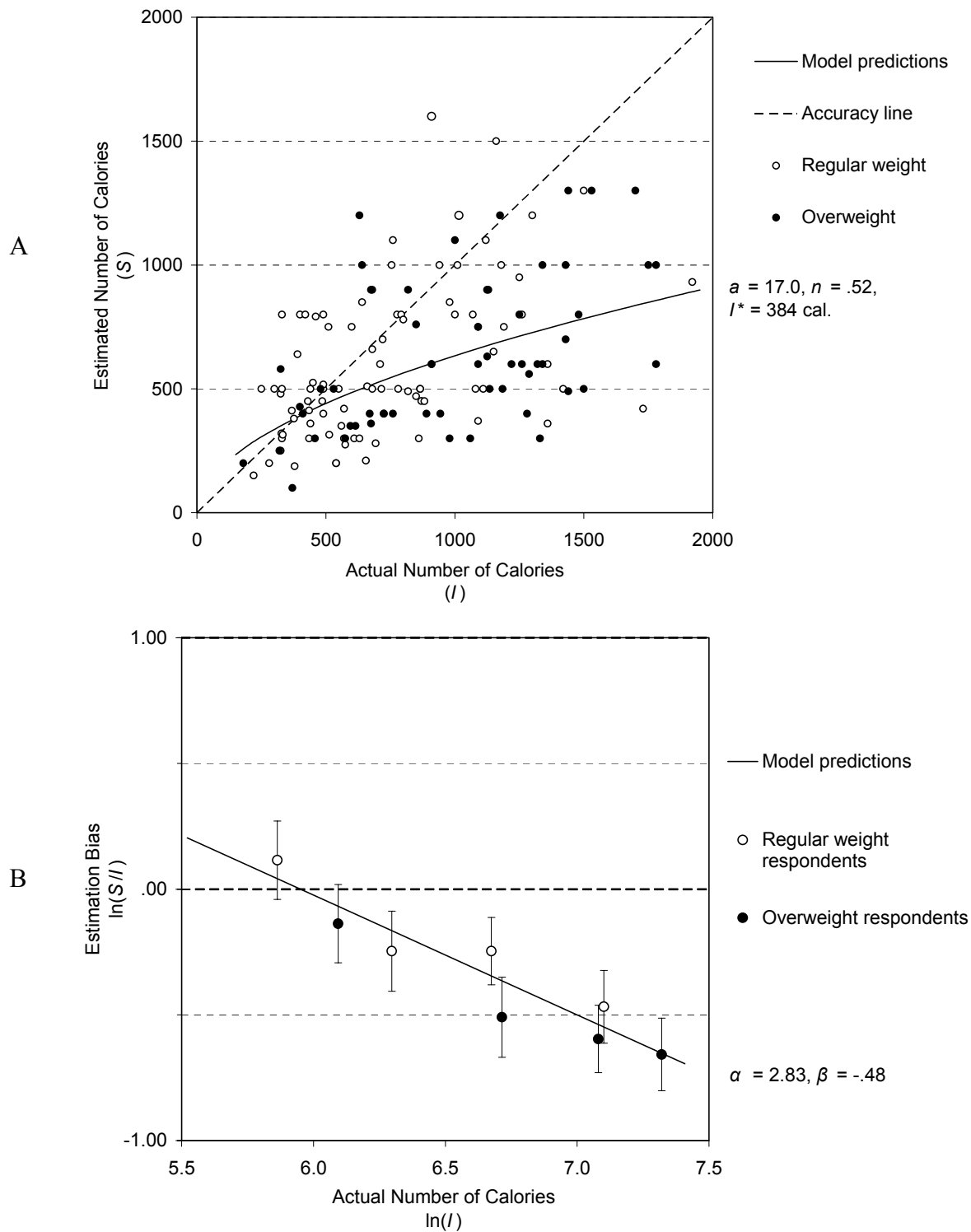


FIGURE 3

EXPERIMENT 2: PREDICTED AND OBSERVED (MEAN AND 95% CI) ESTIMATION BIASES AS A FUNCTION OF ACTUAL NUMBER OF CALORIES AND ESTIMATION CARELESSNESS

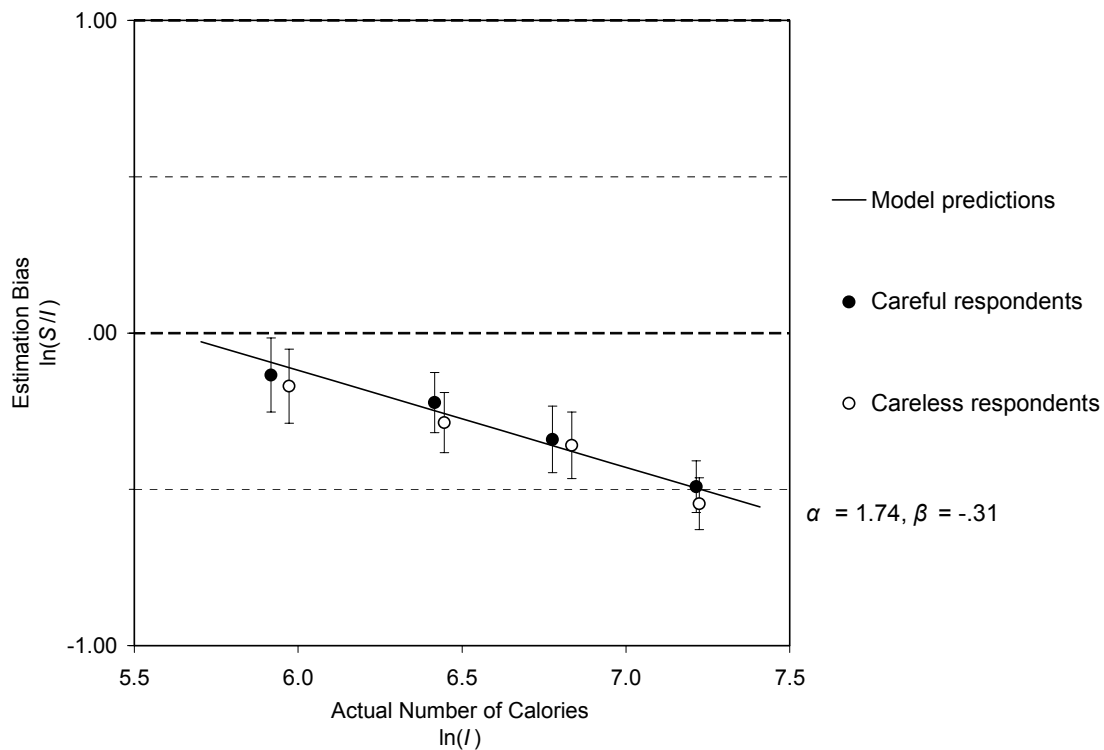


FIGURE 4

EXPERIMENT 3: PREDICTED AND OBSERVED (MEAN AND 95% CI) ESTIMATION BIASES AS A FUNCTION OF ACTUAL NUMBER OF CALORIES AND DEBIASING KNOWLEDGE

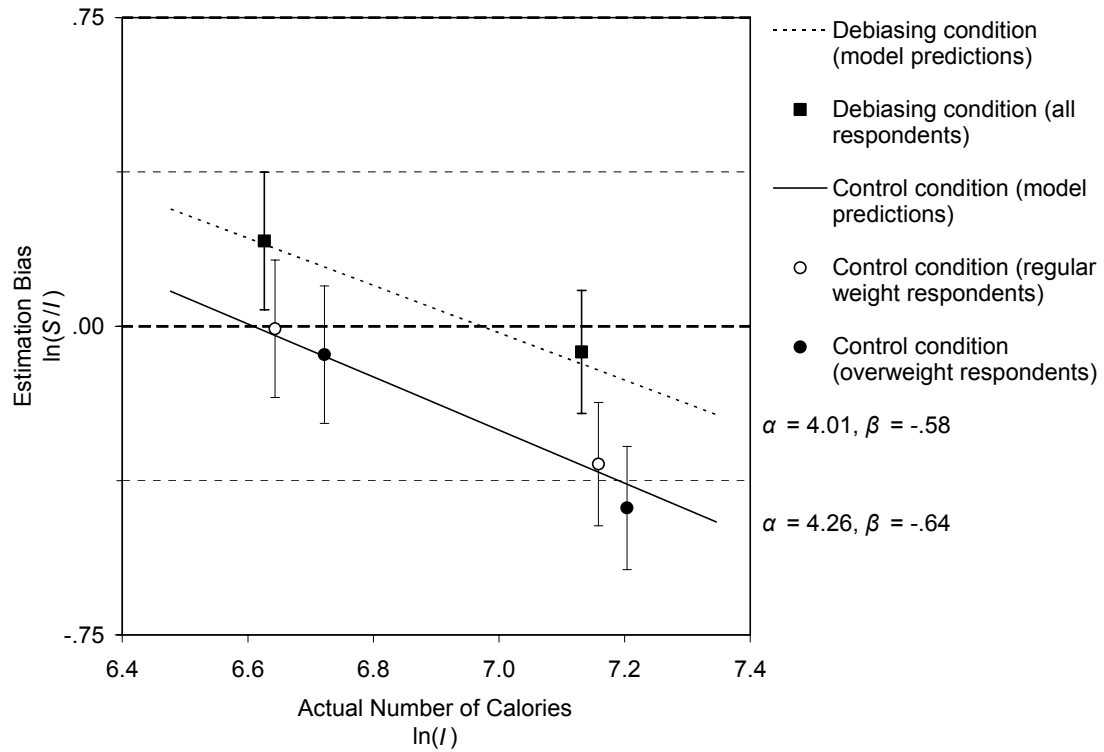


FIGURE 5

EXPERIMENT 4: PREDICTED AND OBSERVED (MEAN AND 95% CI) ESTIMATION BIASES AS A FUNCTION OF ACTUAL NUMBER OF CALORIES, BODY MASS (PANEL A), AND SOCIAL DESIRABILITY (PANEL B)

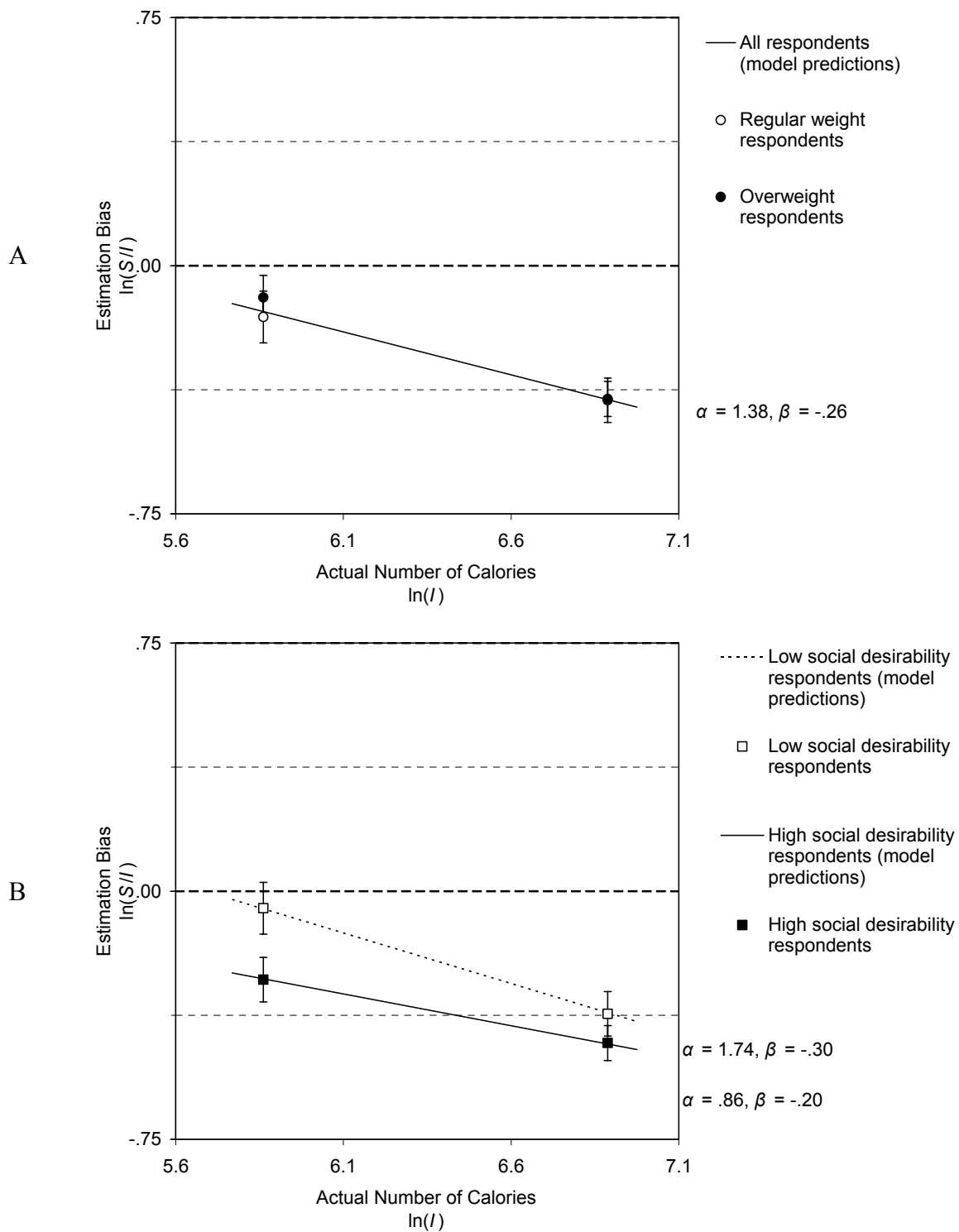


FIGURE 6

EXPERIMENT 5: PREDICTED AND OBSERVED (MEAN AND 95% CI) ESTIMATION BIASES AS A FUNCTION OF ACTUAL NUMBER OF CALORIES AND ESTIMATION ANCHOR

