

How You Estimate Calories Matters: Calorie Estimation Reversals

KAITLIN WOOLLEY

PEGGY J. LIU

Accepted at the Journal of Consumer Research

Kaitlin Woolley (krw67@cornell.edu) is an assistant professor of marketing, Cornell University, Cornell College of Business, 114 East Avenue, Ithaca, NY 14850, USA. Peggy J. Liu (peggy.liu@pitt.edu) is an assistant professor of marketing and Ben L. Fryrear Faculty Fellow, University of Pittsburgh, Katz Graduate School of Business, 3950 Roberto Clemente Drive, Pittsburgh, PA 15260, USA. Please address correspondence to Kaitlin Woolley. The authors thank Jim Bettman, Jeff Inman, Ann L. McGill, Alice Moon, Jay Russo, Manoj Thomas, and Stijn van Osselaer for helpful comments and feedback. The authors also thank Morgan England, Katie Roscoe, Wenxue Zheng and Brad Turner at the Business Simulation Lab at Cornell University and Ed Anderson and the University of Pittsburgh Katz/CBA Business Research Center research assistants for assistance with data collection. Supplementary materials are included in the web appendix accompanying the online version of this article.

OSF Link to code for all analyses, original study materials, and pre-registrations:
<https://osf.io/2emfw>

ABSTRACT

Consumers often form calorie estimates. How consumers estimate calories can systematically bias their calorie assessments. We distinguish between *magnitude estimates*—when consumers judge whether something has “very few” to “many” calories—and *numeric estimates*—when consumers estimate a number of calories. These two estimation modes lead to calorie estimate reversals when assessing calories in stimuli that trade off type and quantity, such as when assessing calories in a smaller portion of unhealthy food versus a larger portion of healthier food. When forming a “magnitude estimate,” people judge the larger, healthier food portion as containing fewer calories than the smaller, unhealthy food portion. However, when forming a “numeric estimate,” people often come to the opposite conclusion—judging the larger, healthier food portion as having more calories. This reversal occurs because these two estimation modes are differentially sensitive to information regarding a stimulus’ type (e.g., food healthiness), which is processed first, and quantity (e.g., food portion size), which is processed secondarily. Specifically, magnitude estimates are more sensitive to type, whereas numeric estimates attend to both type and quantity. Accordingly, this divergence between calorie estimation modes attenuates when: 1) quantity information is made primary or 2) in an intuitive (vs. deliberative) mindset.

Keywords: goal pursuit, calorie estimation, food, exercise, scale compatibility, intuitive versus deliberative processing, portion size

When deciding what to eat, a health-conscious consumer may form an estimate of the calories in different meal options to determine what to select. For example, she may estimate the calories in a smaller, McDonald's cheeseburger and a larger, Subway 12-inch turkey sandwich, and decide to eat the meal she deems has fewer calories.

According to rational theory, different ways in which this consumer forms her calorie estimate *should* lead to the same decision. Specifically, whether she forms a *magnitude estimate*, by judging whether each meal contains somewhere from “very few calories” to “a lot of calories,” versus forms a *numeric estimate*, by estimating the number of calories in each meal, should lead to the same conclusion about which meal is higher in calories. If she judges a smaller McDonald's cheeseburger as having “many calories” and a larger Subway sandwich as having “few calories,” then her numeric estimates should, according to this rational view, reflect the same pattern: a higher estimated number for the smaller cheeseburger (e.g., 500 calories) than the larger turkey sandwich (e.g., 300 calories). However, we propose that forming magnitude versus numeric estimates can actually lead her to *different* conclusions about which meal contains more calories. In particular, we propose that these two estimation modes diverge when evaluating pairs of stimuli that trade-off on “type” and “quantity.” Compared to numeric estimates, we argue that magnitude estimates are more sensitive to information about a stimulus' type (e.g., unhealthy cheeseburger vs. healthy turkey sandwich) and less sensitive to its quantity (e.g., a smaller vs. a larger portion).

Understanding whether and how these two estimation modes—magnitude versus numeric estimates—affect calorie assessments is important for at least three reasons. First, both approaches are commonly used by consumers (see a pilot study in the web appendix). Calorie

estimates—and biases in such estimates—may therefore affect choices aimed at managing caloric intake and expenditure.

Second, there is a disparity in the health advice provided to consumers on how to evaluate their calorie intake: some health advice sources and applications encourage estimating numeric counts in food portions (i.e., form a numeric estimate), as calorie counting has strong associations with weight loss (FDA 2004; Hartmann-Boyce et al. 2014). However, other sources argue that rather than meticulously tracking exact numbers, consumers should consider whether a given food portion is “low” or “high” in calories (i.e., form a calorie magnitude judgment; Borresen 2020; Langer 2017; Wilson 2019).

Third, both estimation modes are used interchangeably in the literature. Whereas some research measures magnitude estimates (e.g., Charbonnier et al. 2016; Slotterback, Leeman, and Oakes 2006), other research measures numeric estimates (Carels, Harper, and Konrad 2006; Chernev 2011), and yet other research measures both (Dohle, Rall, and Siegrist 2014; Liu et al. 2019), treating them as conceptually equivalent (table A1 in the appendix provides examples of prior research). If calorie estimates reverse as a function of estimation mode, this could have implications for conclusions based on prior findings.

THEORETICAL CONCEPTUALIZATION

Type and Quantity as Key Inputs to Calorie Estimates

When estimating calories, two key properties of a stimulus that may affect estimates are

- 1) Type—a categorical judgment of “what” a stimulus is (e.g., a healthy vs. unhealthy food) and
- 2) Quantity—how much of a stimulus there is (e.g., portion size). Prior research suggests that these two properties of stimuli, type and quantity, are processed and coded differently.

Type is considered primary as it is evaluated more quickly and automatically than quantity. Consumers quickly categorize the type of food, for example, by making a judgment about whether the food is healthy or unhealthy (Chernev and Gal 2010; Liu et al. 2019; Oakes 2005; Oakes and Slotterback 2005; Rozin, Ashmore, and Markwith 1996). Because of the primacy of processing food type, type may serve as an initial anchor for subsequent judgments (Liu et al. 2019). Whereas prior research has focused on type in the food domain, similar categorical evaluations based on “type” are likely also made within the exercise domain, with type having primacy of processing there too. For example, exercise workouts can be quickly and automatically judged as relatively “easy” (less intense) or “difficult” (more intense) types of exercise—serving as initial anchors for subsequent judgments.

By contrast, quantity information or “how much of a stimulus there is” is computed more slowly and secondarily relative to type (Liu et al. 2019). That is, when forming judgments, people first consider “what” a stimulus is in terms of its type before considering “how much” of a stimulus there is in terms of its quantity (Liu et al. 2019; Oakes 2005; Ordabayeva and Chandon 2016). Thus, whereas type and quantity can be joint inputs into calories contained in a food portion or calories burned during an exercise workout, these inputs differ in terms of the primacy with which they are processed. Given these key differences, we propose that how type and quantity are incorporated into calorie estimates will differ depending on whether consumers form magnitude versus numeric estimates.

Calorie Estimation Modes are Differentially Sensitive to Type and Quantity

Our focus in this research is in distinguishing between two ways of estimating calories: forming magnitude estimates versus numeric estimates (see table 1 for examples). Objectively, these estimates should move in the same direction: if stimulus A is evaluated as higher in

calories than stimulus B using magnitude estimates, rationally, the same pattern should occur when using numeric estimates.

TABLE 1: DISTINGUISHING BETWEEN TWO CALORIE ESTIMATION MODES

| Calorie Estimate Question (Examples from Different Domains) | Magnitude Estimate | Numeric Estimate |
|---|---|---|
| Food Domain: “How many calories do you think this snack plate has?” | Scale from “very few calories” to “a lot of calories” | Numeric response for number of calories |
| Exercise Domain: “How many calories do you think would be burned by the average female for this activity?” | | |

However, the assumptions of rational theory about the invariance in responses to the method of eliciting responses or preferences may not hold. Indeed, a great deal of research has shown, for example, that whether responses or preferences are assessed with different formats or endpoint labeling can bias judgments in systematic ways (Fischer and Hawkins 1993; Schwarz et al. 1985; Slovic, Griffin, and Tversky 1990; Spence 1990; Thomas and Kyung 2019). Building on this knowledge, we offer the novel proposition that consumers’ calorie estimates diverge as a function of estimation mode under particular circumstances because these estimation modes differentially rely on type versus quantity inputs. Specifically, we suggest that magnitude estimates are more sensitive to type, and less sensitive to quantity, relative to numeric estimates.

Our prediction draws on prior research on scale compatibility—the notion that people are more influenced by attributes that are compatible (vs. incompatible) with the scale on which they are responding (Fischer et al. 1999; Tversky, Sattath, and Slovic 1988). For example, when indicating their preference between gambles, people were more sensitive to a gamble’s payout when providing their willingness to pay (vs. choice), as payment and price are on the same (\$) scale (Tversky et al. 1988). Scale compatibility is typically explained in terms of anchoring and adjustment, wherein people anchor on an attribute that is most compatible with the response mode and then fail to adjust or adjust insufficiently away from it.

Combining the aforementioned literature on scale compatibility with research demonstrating the primacy of type, we suggest that magnitude estimates are more sensitive to type than to quantity. First, we suggest that people forming calorie estimates—regardless of response scale—process type before quantity. Second, and why the estimates differ, is that the response scale when forming a magnitude estimate (few vs. many calories) is more compatible with information on type (i.e., healthy food = fewer calories; less healthy food = more calories). Due to this compatibility, consumers forming magnitude estimates rely on type in providing their response, and are not prompted to incorporate information on quantity. Specifically, after spontaneously thinking about “what” a stimulus is (e.g., a healthy vs. unhealthy type of food), those forming magnitude estimates may fail to adjust much for quantity. By contrast, we reason that numeric estimates engage in more continuous processing. Although information on type is also primary for numeric estimates, type is less compatible with this “number” response scale. This incompatibility prompts those forming numeric estimates to deliberate more and consider other information beyond type, leading them to incorporate quantity into their response.

If magnitude estimates rely on what is primary to a greater extent than numeric estimates, this suggests differences in processing style. That is, we reason that magnitude estimates reflect intuitive, associative thinking, whereas numeric estimates reflect deliberative, calculation-based thinking, in line with prior research (Schley, De Langhe, and Long 2020; Thomas and Morwitz 2009; Viswanathan and Childers 1996; Windschitl and Wells 1996). Dual processing models propose that intuitive processing is automatic and relatively effortless, capturing a gut-reaction, whereas deliberative processing is more effortful and controlled (Chaiken and Trope 1999; Kahneman and Frederick 2002; Sloman 2014). If magnitude estimates anchor on what is primary, they are likely to occur more intuitively than numeric estimates, which are prompted to

incorporate quantity information into their assessment, to the extent that people have time to deliberate.

To summarize, we predict that magnitude versus numeric estimates are differentially sensitive to informational inputs. Magnitude estimates will rely more on type, which is primary (i.e., is a food more or less healthy; is a workout low or high on intensity), and rely less on quantity, which is secondary (i.e., is a food portion smaller or larger; is a workout shorter or longer) compared with numeric estimates. Of importance, our theory predicts when a “calorie estimate reversal” will occur (i.e., that those forming magnitude estimates will judge an unhealthy, smaller food portion as containing more calories than a healthy, larger food portion, which will attenuate or reverse when forming numeric estimates), with implications for understanding how consumers process type and quantity inputs as a function of caloric estimation mode.

When Will Magnitude Estimates and Numeric Estimates Diverge?

If magnitude and numeric estimates are differentially sensitive to type versus quantity, these estimation modes should lead to calorie estimate reversals in specific instances when type and quantity diverge (e.g., for stimuli that trade-off between type and quantity). We thus focus on testing our theory in such situations. For instance, returning to our opening example, suppose that a consumer were deciding between a smaller McDonald’s cheeseburger and a larger Subway 12-inch turkey sandwich. Our theory would predict that if she were forming a magnitude estimate from “very few” to “very many” calories, she would consider the food type (and then not adjust much for quantity, as her food type judgment is more compatible with the magnitude estimate scale). As such, she may estimate that the smaller cheeseburger (vs. 12-inch turkey sandwich) contains more calories. By contrast, if she were forming a numeric estimate (i.e., a

number of calories), she would think of the food type, but because this information is less compatible with a numeric calorie response, she will be prompted to adjust her estimate, thus incorporating information with regard to quantity. Given that food type and quantity diverge in this example, she may even estimate that the smaller cheeseburger (vs. 12-inch turkey sandwich) contains fewer calories when forming a numeric estimate.

Although we focus primarily on calorie estimates of food portions, our theory is not isolated to the food domain. We make analogous predictions for estimating the calories burned in exercise workouts. For example, consider a consumer deciding between a harder, shorter workout versus an easier, longer workout and either estimating calories burned via a magnitude or a numeric estimate. Our theory predicts that magnitude estimates of calories burned will be greater for a harder, shorter workout (vs. easier, longer workout), which will attenuate or reverse when forming numeric estimates of calories burned.

Overall, we predict that magnitude estimates will be more affected by type and less affected by quantity relative to numeric estimates, such that when type and quantity trade off, circumstances arise in which there is an interaction between estimation mode and type/quantity on (standardized) calorie estimates, such that:

- H1a:** Magnitude estimates are lower for healthier, larger food portions (vs. less healthy, smaller portions) and easier, longer workouts (vs. harder, shorter workouts).
- H1b:** Numeric estimates do not exhibit this pattern and may even reverse, such that they are greater for healthier, larger food portions (vs. less healthy, smaller portions) and easier, longer workouts (vs. harder, shorter workouts).

We expect that this pattern occurs because type has primacy of processing and is more compatible with the magnitude estimate response scale. Numeric estimates instead are less compatible with type, prompting individuals to adjust from type information to incorporate

quantity information as well, which requires greater deliberation. We examine this process in two ways.

First, we test for moderation by increasing the primacy of quantity information. If those forming magnitude estimates spontaneously anchor on type and fail to naturally adjust based on quantity (whereas those forming numeric estimates already incorporate quantity), then we anticipate that the divergence of estimation mode as a function of type/quantity will attenuate when quantity information is made more primary. Specifically, when people are first prompted to consider quantity information, magnitude estimates will more closely resemble numeric estimates, suggesting that magnitude estimates rely relatively less on quantity information:

H2: The interaction between estimation mode and type/quantity attenuates when quantity information is made primary. In this case, magnitude estimates incorporate quantity information more, such that they resemble numeric estimates.

Second, we test for moderation by processing style. If numeric estimates anchor on type, which is primary, but adjust based on quantity due to scale incompatibility, then prompting people who are making numeric estimates to increase their speed of processing by adopting an intuitive (vs. deliberative) mindset should reduce their reliance on quantity information. As such, we hypothesize that when prompted to form a quick, gut response, numeric estimates will more closely resemble magnitude estimates, suggesting that numeric estimates first anchor on type and adjust away, relying on deliberation to incorporate quantity into their response.

H3: The interaction between estimation mode and type/quantity attenuates when in an intuitive (vs. deliberative) mindset. When in an intuitive mindset, numeric estimates incorporate quantity information less, such that they resemble magnitude estimates.

PRESENT RESEARCH

Six experiments (five pre-registered) test our hypotheses. Experiment 1 examined estimates of calories in food portions using a paradigm comparing healthier, larger portions with less healthy, smaller portions. Supporting hypothesis 1, consumers estimated fewer calories in a healthier, larger food portion (vs. less healthy, smaller food portion) when forming magnitude estimates, which attenuated when forming numeric estimates, with implications for real food choice. Experiment 2 examined estimates of calories burned from exercise using a paradigm comparing easier, longer exercises with harder, shorter exercises, holding total calories burned equivalent. People estimated fewer calories burned for easier, longer exercises (vs. harder, shorter exercises) when forming magnitude estimates, which reversed when forming numeric estimates.

Experiments 3-6 returned to focusing on assessing calories in portions of food. Experiment 3 replicated the key effect of estimation mode and demonstrated that estimation mode moderates findings from prior research. Specifically, previous results asking consumers to form numeric estimates in sandwiches of different sizes from two popular chain restaurants (McDonald's cheeseburger vs. Subway 12-inch turkey sandwich) replicated when consumers formed numeric estimates as in the original study (Chandon and Wansink 2007, study 2), but reversed when consumers formed magnitude estimates.

Experiments 4-6 provided evidence for our proposed process. Experiment 4 fully tested our overarching premise for hypothesis 1 by fully crossing food type (more vs. less healthy) and quantity (high vs. low), demonstrating that magnitude (vs. numeric) estimates were more sensitive to differences in food type, and less sensitive to differences in portion size. Examining the process further, experiment 5 provided evidence for hypothesis 2, showing that increasing the primacy of quantity information attenuates the divergence between these two estimation modes,

by making magnitude estimates mirror numeric estimates. Lastly, experiment 6 provided evidence for hypothesis 3, showing that adopting an intuitive (vs. deliberative) mindset also attenuates the divergence between these two estimation modes, this time by making numeric estimates mirror magnitude estimates.

Across experiments, we report all manipulations, all measures, and all exclusions (see the web appendix for exclusion criteria). We pre-registered five experiments (experiments 2-6) at AsPredicted.org. Data, syntax, materials, and pre-registrations are available at OSF: <https://osf.io/2emfw>. In terms of our analysis approach, for most experiments, we first analyzed each estimation mode separately, as calorie estimates were on different scales across estimation mode conditions (e.g., testing whether the healthier, larger food portion was rated as having more or fewer calories than the less healthy, smaller food portion within each estimation mode). We then computed a standardized z -score to put calorie estimates on a comparable scale across estimation mode conditions, and examined the interaction between food type/quantity and estimation mode. Experiment 1 also reports calorie estimate reversals and includes a table that summarizes such reversals across all of the experiments that assessed calorie estimates for stimuli trading off type/quantity within-person (experiments 1, 3, 5, and 6).

EXPERIMENT 1: REVERSAL IN MAGNITUDE VERSUS NUMERIC ESTIMATES OF CALORIES IN FOOD

Experiment 1 examined our prediction that magnitude versus numeric estimates diverge when type/quantity trade-off in the food domain. We assessed calorie estimates for two snack plates (a less healthy, smaller food portion and a healthier, larger food portion) that differed in their actual caloric content. The reason for this set-up was twofold. First, we measured actual

food choice after assessing calorie estimates, instructing participants to choose a snack plate they thought contained fewer calories to eat in the lab. Second, we tested whether divergence in calorie estimates can cause people to make a mistake—at times leading them to choose a snack plate containing more calories. We predicted that when forming a magnitude estimate (judgment from few to many calories), people would estimate fewer calories in a healthier, larger snack plate than a less healthy, smaller snack plate, but that when forming a numeric estimate (a number of calories), this pattern would not exist and might even reverse. We also expected estimation mode to affect choice, such that more people forming magnitude (vs. numeric) estimates would choose the healthier, larger snack plate (200 calories) over the less healthy, smaller snack plate (100 calories), which would be mediated by differences in calorie estimates for the two snack plates.

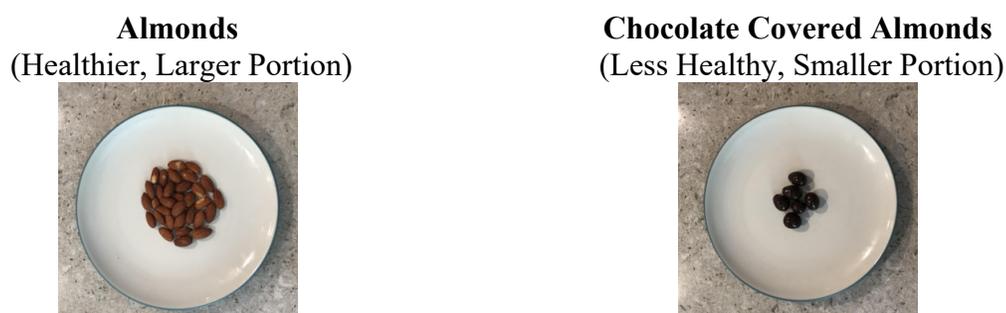
Method

A total of 276 undergraduates at a U.S. university completed this experiment in the lab for course credit. We excluded participants with allergies to the snack options and participants with numeric estimates 3 SD above the mean, leaving 243 participants ($M_{\text{age}} = 19.70$, $SD = 1.31$; 53.1% female; $M_{\text{BMI}} = 23.23$) ($n = 120$ per between-subjects cell). For all experiments, we test for effects and interactions with age, gender, and BMI (reported in the web appendix).

Participants were randomly assigned to condition in a 2 (estimation mode: magnitude vs. numeric estimate; between-subjects) \times 2 (type/quantity: healthier, larger vs. less healthy, smaller; within-subject) mixed-model design. Participants viewed a photo of a snack plate of a larger portion of almonds (33g portion size; 200 calories, although participants only saw the photo and did not receive this calorie information) and a snack plate of a smaller portion of chocolate

covered almonds (20g; 100 calories) (order counterbalanced; see figure 1). In the magnitude estimate condition, for each snack plate, we asked, “How many calories do you think this snack has?” (1 = “very few calories” to 9 = “a lot of calories”). In the numeric estimate condition, we asked the same question as an open numeric response.

FIGURE 1: SNACK PLATE AS A FUNCTION OF TYPE/QUANTITY (EXPERIMENT 1)



Participants viewed images of each snack plate again and were instructed to choose a low-calorie option for their snack. Participants selected their chosen snack from one of two brown paper bags located at their desk (see the web appendix for the lab set up with photos). Participants ate their snack while listening to an 11-minute audio program. We included manipulation checks to assess perceived healthiness and quantity for each snack plate (1 = “not at all healthy/very small” to 9 = “very healthy/very large”).

Results

Manipulation checks. We confirmed that participants perceived the larger portion of almonds (vs. smaller portion of chocolate covered almonds) as healthier ($M_{\text{almonds}} = 7.16$; $M_{\text{chocolate covered almonds}} = 4.56$; $t(242) = 23.51, p < .001, d = 1.51$) and as a larger portion ($M_{\text{almonds}} = 5.58$; $M_{\text{chocolate covered almonds}} = 2.85$; $t(242) = 28.31, p < .001, d = 1.82$).

Magnitude estimates. As predicted, participants forming magnitude estimates estimated significantly fewer calories in a larger portion of almonds than in a smaller portion of chocolate covered almonds ($M_{\text{almonds}} = 4.03$, $SD = 1.97$; $M_{\text{chocolate covered almonds}} = 5.24$, $SD = 1.70$; $t(121) = 6.50$, $p < .001$, $d = .59$; non-parametric Wilcoxon test: $z = 5.70$, $p < .001$).

Numeric estimates. In contrast to the magnitude estimate results, there was no significant difference in numeric estimates as a function of type/quantity ($M_{\text{almonds}} = 116.83$, $SD = 77.94$; $M_{\text{chocolate covered almonds}} = 110.70$, $SD = 59.02$; $t(120) = .99$, $p = .325$, $d = .09$; non-parametric Wilcoxon test: $z = .56$, $p = .577$).

Interaction between estimation mode and type/quantity. A repeated measures ANOVA of standardized calorie estimates on estimation mode \times type/quantity yielded a significant interaction ($F(1, 241) = 29.32$, $p < .001$, $\eta_p^2 = .11$), indicating that magnitude (vs. numeric estimates) diverged as a function of type/quantity.

Calorie estimate reversals. This experiment measured calorie estimates for the different food stimuli within-subject, allowing us to examine whether estimation mode leads to reversals in calorie estimates. A chi-square analysis revealed that significantly more people estimated a larger portion of almonds as having fewer calories than a smaller portion of chocolate covered almonds when forming magnitude estimates (79.2%) than numeric estimates (52.7%), $\chi^2(1, N = 192) = 15.08$, $p < .001$, $\phi = .28$). This analysis excludes participants indicating identical responses for both items, following preference reversal research (e.g., Slovic et al. 1990; O'Donnell and Evers 2019), which did not significantly differ by condition (magnitude estimate: 17.2%; numeric estimate: 24.8%; $\chi^2(1, N = 243) = 2.10$, $p = .147$, $\phi = -.09$).

Table 2 summarizes analyses of calorie estimate reversals for all experiments in which we assessed stimuli diverging on type/quantity within-subject (i.e., experiments 1, 3, 5, and 6).

Following prior preference reversals analyses that measured indifferent responses across conditions and excluded indifferent responses (O'Donnell and Evers 2019), we excluded participants indicating identical responses for both items in these analyses. Results of these calorie estimate reversal analyses generally converge with the conclusions of the calorie estimate analyses reported earlier (and which are applicable across the designs of all experiments). We refer readers to table 2 for an overall summary of the calorie estimate reversals across studies in which we assessed stimuli diverging on type/quantity within-subject and report full details of these analyses in the web appendix. The web appendix also contains a table examining the effect of estimation mode when categorizing responses into one of three options (i.e., healthier, larger snack plate has fewer, more, or equal calories as the less healthy, smaller snack plate).

TABLE 2: SUMMARY OF CALORIE ESTIMATE REVERSALS FOR EXPERIMENTS ASSESSING ESTIMATES OF TYPE/QUANTITY WITHIN-SUBJECT

| Exp | N | Manipulation | Healthier, larger snack plate has fewer calories | | Significance Test ^a |
|-----|-----------------------|-------------------------|---|----------------------------------|--|
| | | | Magnitude Estimate ^a | Numeric Estimate ^a | |
| 1 | 243 (51 identical) | | 79.2% (80/101) | 52.7% (48/91) | $\chi^2(1, N = 192) = 15.08, p < .001, \phi = .28$ |
| 3 | 277 (25 identical) | | 79.8% (95/119) | 44.4% (59/133) | $\chi^2(1, N = 252) = 33.25, p < .001, \phi = .36$ |
| 5 | 397 (52 identical) | Control | 58.8% (100/170) | 42.3% (74/175) | $\chi^2(1, N = 345) = 9.43, p = .002, \phi = .17$ |
| | 391 (41 identical) | Quantity Primary | 38.4% (66/172) | 34.8% (62/178) | $\chi^2(1, N = 350) = .47, p = .492, \phi = .04$ |
| 6 | 476 (44 identical) | Deliberative Mindset | 65.0% (132/203) | 49.3% (113/229) | $\chi^2(1, N = 432) = 10.78, p = .001, \phi = .16$ |
| | 466 (31 identical) | Intuitive Mindset | 65.5% (129/197) | 57.1% (136/238) | $\chi^2(1, N = 435) = 3.15, p = .076, \phi = .09$ |

Notes. ^aThese columns exclude participants who provided identical calorie estimates for the two food items diverging on type/quantity. Details for the calorie reversal analyses for experiments 3, 5, and 6, and analyses including all participants, are reported in the web appendix.

Food choice. Estimation mode significantly affected food choice; more participants chose to eat a larger portion of almonds when forming magnitude (vs. numeric estimates) (59.8% vs. 46.3%; $\chi^2(1, N = 243) = 4.48, p = .034, \phi = .14$). The effect of estimation mode on food choice

was mediated by the difference between standardized calorie estimates of the two snack plates (chocolate-covered almonds minus almonds; $\beta_{\text{index}} = 1.46$, $SE = .42$, $95\% \text{ CI} = [.85, 2.47]$).

Discussion

Experiment 1 supported our hypothesis that magnitude and numeric estimates diverge as a function of type/quantity. When forming magnitude estimates, consumers estimated fewer calories in a healthier, larger food portion (vs. a less healthy, smaller food portion; hypothesis 1a), which attenuated when forming numeric estimates (hypothesis 1b). We further find a significant calorie estimate reversal, such that more participants estimated the larger portion of almonds as containing fewer calories than the smaller portion of chocolate covered almonds when forming magnitude (vs. numeric) estimates.

This experiment further demonstrated a real consequence of calorie estimation mode for food choice. When instructed to choose the lower calorie snack plate, participants were more likely to choose the larger portion of almonds (vs. the smaller portion of chocolate covered almonds) when forming magnitude (vs. numeric) estimates, even though the larger portion of almonds actually contained more calories. This was of consequence to participants, who were provided with and instructed to eat their selected snack in the lab.

EXPERIMENT 2: REVERSAL IN MAGNITUDE VERSUS NUMERIC ESTIMATES OF CALORIES BURNED FOR DIFFERENT EXERCISES

Experiment 2 tested our prediction within the exercise domain, examining a reversal in magnitude and numeric estimates of calories burned for easier, longer workouts versus harder, shorter workouts. In addition, each participant provided both magnitude and numeric estimates of calories burned, such that estimation mode was a within-subject factor.

To test the robustness of the estimate reversal observed in experiment 1, we selected exercise intensity and quantity levels to equate exercises in terms of calories burned across type/quantity (e.g., 30 minutes of slow biking and 10 minutes of moderate biking both burn approximately 100 calories for the average 150 lb. female). We anticipated that magnitude estimates would be lower for easier, longer exercises than for harder, shorter exercises, which would reverse for numeric estimates.

Method

We pre-registered this experiment and recruited 403 participants from Prolific ($n = 200$ per between-subjects cell). We randomly assigned participants to condition in a 2 (estimation mode: numeric estimate vs. magnitude estimate; within-subject) \times 2 (intensity/quantity: easier, longer vs. harder, shorter; between-subjects) \times 3 (exercise replicate: biking vs. yoga vs. treadmill; within-subject) mixed model design. As pre-registered, we excluded participants with numeric estimates 3 SD above the mean, leaving 397 participants ($M_{\text{age}} = 33.45$, $SD = 11.24$; 56.2% female; $M_{\text{BMI}} = 27.32$).

Participants evaluated three exercises corresponding to three calorie replicates that an average 150 lb. female engaged in on three different recent days (table 3). For easier, longer workouts, participants viewed slow biking, low intensity yoga, and slow treadmill pace for either 15, 30, or 45 minutes (activity and quantity replicates were randomly assigned and counterbalanced). For harder, shorter workouts, participants viewed moderate intensity biking, high intensity yoga, and fast treadmill pace for either 5, 10, or 15 minutes. Exercises burned similar calories across intensity/quantity condition.

TABLE 3: EXERCISE INTENSITY/QUANTITY FOR TOTAL CALORIE REPLICATES FOR AN AVERAGE (150 LB.) FEMALE

| | 50-calorie replicate | 100-calorie replicate | 150-calorie replicate |
|---|----------------------|-----------------------|-----------------------|
| Less intense, longer stimuli | | | |
| Slow biking (low intensity: < 10mph) | 15 minutes | 30 minutes | 45 minutes |
| Yoga (low intensity: stretching/hatha yoga) | 15 minutes | 30 minutes | 45 minutes |
| Slow treadmill pace (low intensity: 3.2 mph) | 15 minutes | 30 minutes | 45 minutes |
| More intense, shorter stimuli | | | |
| Biking (moderate intensity: 14mph-15.9mph) | 5 minutes | 10 minutes | 15 minutes |
| Yoga (high intensity: power yoga) | 5 minutes | 10 minutes | 15 minutes |
| Fast treadmill pace (moderate intensity: 5.2 mph) | 5 minutes | 10 minutes | 15 minutes |

In the magnitude estimate condition, we asked “How many calories do you think would be burned by the average female for this activity?” (1 = “very few calories” to 9 = “a lot of calories”). In the numeric estimate condition, we asked the same question as an open numeric response. We blocked measures together by estimation mode (block order counterbalanced). For example, participants provided magnitude (*numeric*) estimates for three exercises and then numeric (*magnitude*) estimates for the same three exercises. We included manipulation checks to assess quantity and intensity (1 = “not very long/intense” to 9 = “very long/intense”).

Results

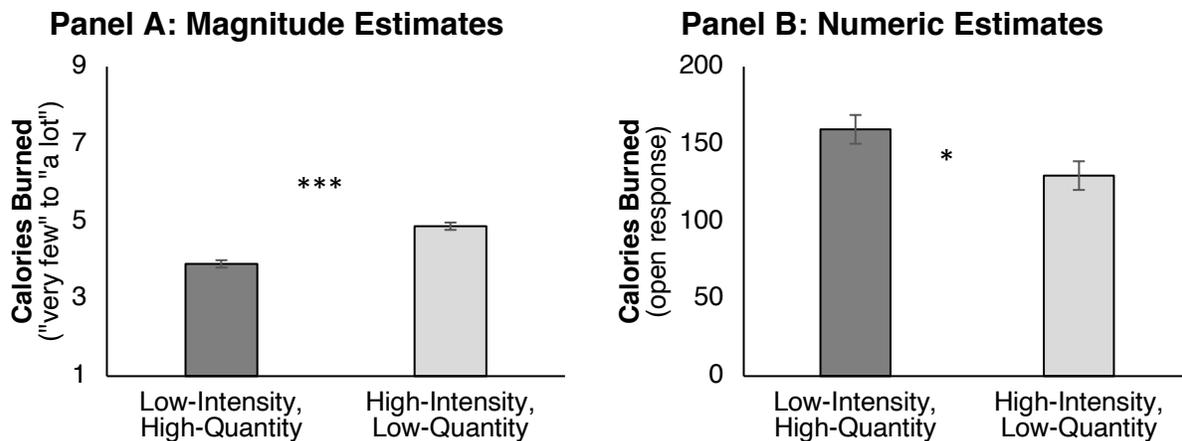
Manipulation checks. We first confirmed our manipulation of exercise intensity ($M_{\text{low-intensity, longer}} = 2.92$; $M_{\text{high-intensity, shorter}} = 5.48$; $F(1, 395) = 501.12, p < .001, \eta_p^2 = .56$) and quantity ($M_{\text{low-intensity longer}} = 4.43$; $M_{\text{high-intensity, shorter}} = 3.70$; $F(1, 395) = 31.80, p < .001, \eta_p^2 = .07$).

Magnitude estimates of calories burned. A repeated measures ANOVA of magnitude estimates on intensity/quantity for the three exercise replicates revealed that participants expected to burn significantly fewer calories when forming magnitude estimates of calories burned for easier, longer workouts ($M = 3.90, SD = 1.17$) than harder, shorter workouts ($M = 4.88, SD = 1.48$; $F(1, 395) = 52.75, p < .001, \eta_p^2 = .12$; non-parametric Kruskal-Wallis test: $\chi^2 =$

47.77, $p < .001$; figure 2A). There was a main effect of exercise replicate ($F(2, 394) = 38.06$, $p < .001$, $\eta_p^2 = .16$) and no significant interaction ($F(2, 394) = .71$, $p = .495$, $\eta_p^2 < .01$). We find the same effect when using calorie replicates, with no significant interaction ($F < 1$).

Numeric estimates of calories burned. Opposite to the magnitude estimate results, participants forming numeric estimates estimated burning more calories for easier, longer workouts ($M = 159.48$, $SD = 129.86$) than harder, shorter workouts ($M = 129.59$, $SD = 131.74$; $F(1, 395) = 5.18$, $p = .023$, $\eta_p^2 = .01$; non-parametric test: $\chi^2 = 15.56$, $p < .001$; figure 2B). There was a main effect of exercise replicate ($F(2, 394) = 15.40$, $p < .001$, $\eta_p^2 = .07$), and no significant interaction ($F(2, 394) = 2.01$, $p = .136$, $\eta_p^2 = .01$). We find the same effect when using calorie replicates, with no significant interaction ($F < 1$).

FIGURE 2: DIVERGENCE IN MAGNITUDE (VS. NUMERIC) ESTIMATES OF CALORIES BURNED AS A FUNCTION OF ESTIMATION MODE (EXPERIMENT 2)



Note. *** $p < .001$. Bars indicate ± 1 SEM.

Interaction between estimation mode and intensity/quantity. To examine the interaction between estimation mode and exercise intensity/quantity, we standardized magnitude and numeric estimates of calories burned, in line with our pre-registration. A repeated measures

ANOVA of calorie assessment on estimation mode, intensity/quantity, and their interaction for the three exercise replicates resulted in the predicted estimation mode \times intensity/quantity interaction ($F(1, 395) = 56.52, p < .001, \eta_p^2 = .13$). There was a main effect of intensity/quantity ($F(1, 395) = 7.49, p = .006, \eta_p^2 = .02$) with no other significant effects or interactions.¹

Discussion

Experiment 2 demonstrated that the effects observed in experiment 1 extend beyond estimating calories in food portions to estimating calories burned during a workout. Supporting our hypothesis that magnitude versus numeric estimates diverge as a function of intensity/quantity, people forming magnitude estimates anticipated that easier, longer exercises would burn fewer calories than harder, shorter exercises (hypothesis 1a), and anticipated the opposite when forming numeric estimates (hypothesis 1b). Thus, unlike experiment 1 which demonstrated attenuation of the type/quantity effect within numeric estimates, in experiment 2 we observed a significant reversal of the type/quantity effect within numeric estimates (relative to magnitude estimates). We return to this point on why numeric estimates may at times attenuate and other times reverse in the General Discussion.

Finally, by showing this effect using calorie estimation mode as a within-subject factor (i.e., joint evaluation mode), this experiment also addresses a potential alternative explanation based on joint versus separate evaluation mode (Hsee 1996), wherein reversals might occur because magnitude estimates are less evaluable, and thus quantity insensitive, in separate

¹ This analysis standardized and analyzed values within exercise replicate (yoga vs. biking vs. treadmill). We find a similar interaction when standardizing and analyzing within total calorie replicate ($F(1, 395) = 57.49, p < .001, \eta_p^2 = .13$; figures S1 and S2 in the web appendix).

evaluation. Having found that this calorie estimate reversal generalizes to the exercise domain, our remaining experiments return to the food domain.

EXPERIMENT 3: REPLICATION & MODERATION OF PRIOR FINDINGS

Experiment 3 examined a consequence of the divergence between magnitude estimates and numeric estimates for consumer research by testing for a replication and moderation of an existing finding in the literature. As noted in table A1 in the appendix, research typically assesses calories using a single approach (i.e., either magnitude or numeric estimates). Even when both estimation measures are assessed in the same study, research typically treats these as interchangeable measures of the same calories construct. Sometimes, this approach may produce similar conclusions: that is, magnitude and numeric estimates may operate in the same direction. However, we have proposed and shown that these estimation modes can systematically diverge when type and quantity trade-off.

To further demonstrate the importance and relevance of this finding, experiment 3 aimed to replicate and then test for moderation of an effect observed in prior research, wherein people forming numeric estimates estimated more calories in a healthier, larger meal (Subway 12-inch turkey sandwich) than in a less healthy, smaller meal (McDonald's cheeseburger; study 2 in Chandon and Wansink 2007). We tested for a pre-registered replication of two of the conditions from study 2 in C&W 2007 *JCR* and then tested for our predicted reversal in two added conditions. We asked participants to form either magnitude or numeric estimates of calories in these two food items, predicting that when people form numeric estimates, as in the original study, we would replicate this effect, but that it would reverse when people formed magnitude estimates.

Method

We pre-registered this experiment and recruited 300 participants from MTurk ($n = 150$ per between-subjects cell). As in the original experiment, to qualify to participate, participants needed to have eaten at McDonald's and Subway at least three times in the past year. We carried out exclusions in line with our pre-registration (see web appendix), for a final sample of 277 ($M_{age} = 39.18$, $SD = 11.93$; 52.3% female; $M_{BMI} = 28.28$).

Participants were randomly assigned to condition in a 2 (estimation mode: magnitude vs. numeric estimate; between-subjects) \times 2 (type/quantity: healthier, larger vs. less healthy, smaller; within-subject) mixed-model design. All participants were asked to assess calories in two food items (counterbalanced): a Subway 12-inch turkey sandwich (healthier, larger; 510 calories) and a McDonald's cheeseburger (less healthy, smaller; 300 calories). These calorie amounts were taken from nutrition information reported on each restaurant's website in 2019 and were not displayed to participants. In the magnitude estimate condition, we asked participants, "How many calories do you think a Subway 12-inch turkey sandwich [*McDonald's cheeseburger*] contains?" on a nine-point scale from "very few calories" to "a lot of calories." In the numeric estimate condition, we asked the same question as an open numeric response.

As manipulation checks, we measured perceived healthiness and size of each food item on nine-point scales anchored at "not at all healthy/very small" to "very healthy/very large." We also included a five-item nine-point nutrition involvement scale ($\alpha = .92$), as in the original experiment, for exploratory purposes. Additionally, we asked participants whether they looked up information about how many calories the food items contained (yes/no), which we used to carry out pre-registered exclusions.

Results

Manipulation checks. A 12-inch Subway sandwich was seen as healthier ($M_{\text{Subway}} = 5.92$; $M_{\text{McDonald's}} = 2.69$; $t(276) = 24.35$, $p < .001$, $d = 1.46$) and larger ($M_{\text{Subway}} = 7.34$; $M_{\text{McDonald's}} = 3.71$; $t(276) = 28.28$, $p < .001$, $d = 1.70$) than a McDonald's cheeseburger.

Calorie estimates. Replicating the pattern from C&W 2007 JCR, numeric estimates of calories in a Subway 12-inch turkey sandwich were greater than numeric estimates of a McDonald's cheeseburger ($M_{\text{Subway}} = 545.37$, $SD = 251.89$; $M_{\text{McDonald's}} = 499.92$, $SD = 259.71$; $t(136) = 1.90$, $p = .059$, $d = .16$; non-parametric Wilcoxon test: $z = -2.28$, $p = .023$). However, this pattern reversed when assessing magnitude estimates; this time, people estimated significantly fewer calories in a Subway sandwich than in a McDonald's cheeseburger ($M_{\text{Subway}} = 5.81$, $SD = 1.58$; $M_{\text{McDonald's}} = 7.31$, $SD = 1.58$; $t(139) = -9.27$, $p < .001$, $d = -.78$; non-parametric Wilcoxon test: $z = 7.38$, $p < .001$).

Interaction between estimation mode and type/quantity. A repeated measures ANOVA on standardized calorie estimates resulted in a significant interaction ($F(1, 275) = 62.25$, $p < .001$, $\eta_p^2 = .18$), indicating estimates significantly diverged as a function of type/quantity.

Moderation by nutrition involvement. We also conducted exploratory analyses on whether nutrition involvement ($M = 6.17$, $SD = 2.05$, median = 6.60) differentially affected magnitude estimates versus numeric estimates. In the original experiment, participants forming numeric estimates with greater (vs. lower) nutrition involvement were more accurate in that they were more likely to estimate that a Subway sandwich had more calories than a McDonald's cheeseburger (table 4). We replicated this result when assessing numeric estimates ($F(1, 135) = 8.09$, $p = .005$, $\eta_p^2 = .06$). By contrast, this corresponding interaction was not significant when assessing magnitude estimates ($F(1, 138) = .47$, $p = .493$, $\eta_p^2 < .01$). These analyses relied on

raw estimates, in keeping with the original paper (Chandon and Wansink 2007). Examining the interaction between estimation mode, food type/quantity, and nutrition involvement using standardized calorie estimates revealed a non-significant three-way interaction ($F(1, 273) = 2.11$, $p = .148$, $\eta_p^2 = .01$). These exploratory findings by nutrition involvement are thus further suggestive (though not conclusive, given that the three-way interaction was not significant) that research conclusions about calorie assessments can vary as a function of caloric estimation mode.

TABLE 4: CALORIE ESTIMATES BY NUTRITION INVOLVEMENT

| <i>Numeric estimates (C&W 2007 JCR, study 2)</i> | | |
|--|---------------------------|----------------------------|
| | Low Nutrition Involvement | High Nutrition Involvement |
| McDonald's cheeseburger | 324 | 345 |
| Subway 12-inch turkey sandwich | 353 | 442 |
| <i>Numeric estimates (current paper, experiment 3)</i> | | |
| McDonald's cheeseburger | 543 | 466 |
| Subway 12-inch turkey sandwich | 526 | 560 |
| <i>Magnitude estimates (current paper, experiment 3)</i> | | |
| McDonald's cheeseburger | 7.47 | 7.13 |
| Subway 12-inch turkey sandwich | 5.97 | 5.63 |

Note. Analysis treats nutrition involvement on a continuous scale; numbers in table are values at median split for nutrition involvement.

Discussion

This experiment conceptually replicated the divergence in magnitude versus numeric estimates as a function of type/quantity observed in experiments 1-2. This experiment also demonstrated consequences of this divergence for conclusions from prior marketing research. Specifically, we replicated a previous pattern in the literature when consumers formed numeric estimates and then showed that this pattern reversed when consumers formed magnitude estimates. Having documented our key interaction effect across experiments 1-3, our remaining experiments examine the proposed underlying process.

EXPERIMENT 4: MAGNITUDE (VS. NUMERIC) ESTIMATES ARE MORE SENSITIVE TO TYPE THAN TO QUANTITY

Experiment 4 examined our prediction that magnitude estimates are more sensitive to differences in type (i.e., food healthiness), and less sensitive to differences in quantity (i.e., portion size), whereas numeric estimates incorporate both type and quantity such that they are relatively less sensitive to differences in type, and more sensitive to differences in quantity. Using stimuli from experiment 1, we fully crossed food type (almonds vs. chocolate covered almonds) with quantity (portions corresponding to 100 vs. 200 calories). We expected a type \times quantity \times estimation mode interaction, such that magnitude (vs. numeric) estimates would be more sensitive to food type, and less sensitive to quantity.

Method

We pre-registered this experiment and recruited 600 participants from MTurk ($n = 75$ per between-subjects cell). As pre-registered, we excluded those with numeric estimates 3 SD above the mean, leaving 594 ($M_{\text{age}} = 37.16$, $SD = 11.20$; 48.5% female; $M_{\text{BMI}} = 27.21$).

Participants were randomly assigned to one of eight conditions in a 2 (estimation mode: magnitude vs. numeric estimate) \times 2 (type: healthier vs. less healthy) \times 2 (quantity: smaller vs. larger) between-subjects design. Depending on condition, participants viewed a photo of one of the following snack plates (see the web appendix): healthier, smaller – 16.5g almonds (100 calories); healthier, larger – 33g almonds (200 calories); less healthy, smaller – 20g chocolate covered almonds (100 calories); or less healthy, larger – 40g chocolate covered almonds (200 calories).

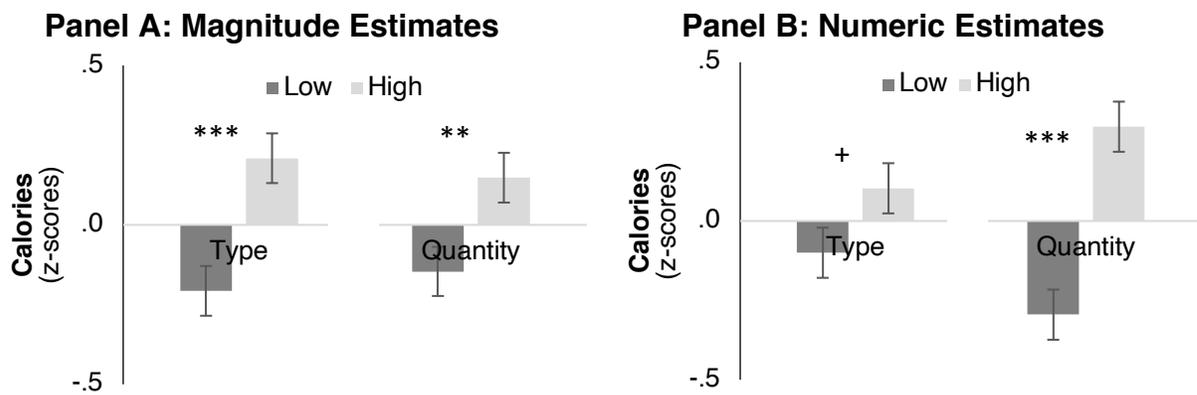
In the magnitude estimate condition, we asked participants, “How many calories do you think this snack has?” (nine-point scale anchored at “very few calories” to “a lot of calories”). In

the numeric estimate condition, we asked, “How many calories do you estimate are in this snack? Please enter a number:” (numeric open response).

Results

To examine whether magnitude estimates are more sensitive to type and less sensitive to quantity than numeric estimates, we tested for a three-way interaction between estimation mode, food type, and quantity, on standardized calorie estimates. An ANOVA of calorie estimates resulted in a significant three-way interaction ($F(1, 586) = 9.31, p = .002, \eta_p^2 = .02$). As predicted, magnitude estimates were significantly more sensitive to differences in food type (i.e., healthiness; $F(1, 586) = 14.36, p < .001, \eta_p^2 = .02$; non-parametric Kruskal-Wallis test: $\chi^2 = 13.78, p < .001$) than quantity ($F(1, 586) = 7.15, p = .008, \eta_p^2 = .01$; non-parametric Kruskal-Wallis test: $\chi^2 = 7.04, p = .008$; figure 3A). However, numeric estimates were less sensitive to differences in type ($F(1, 586) = 3.29, p = .070, \eta_p^2 < .01$; non-parametric Kruskal-Wallis test: $\chi^2 = 1.32, p = .250$) than quantity ($F(1, 586) = 28.12, p < .001, \eta_p^2 = .05$; non-parametric Kruskal-Wallis test: $\chi^2 = 22.19, p < .001$; figure 3B).

FIGURE 3: SENSITIVITY TO TYPE AND QUANTITY WHEN FORMING STANDARDIZED MAGNITUDE ESTIMATES (PANEL A) AND NUMERIC ESTIMATES (PANEL B) (EXPERIMENT 4)



Note. Bars indicate ± 1 SEM; + $p < .10$; ** $p < .01$; *** $p < .001$. “Low” refers to healthier for Type and smaller portion for Quantity; “High” refers to less healthy for Type and larger portion for Quantity.

Discussion

Experiment 4 demonstrated convergent support for our theory, this time when type and quantity were fully crossed. That is, unlike experiments 1-3, this experiment was not designed to test for divergence between estimation modes, as food healthiness was orthogonal to portion size. Given this, we were able to directly examine the sensitivity of magnitude estimates to type (vs. quantity) compared with numeric estimates. Doing so, we found evidence that magnitude estimates are relatively more sensitive to healthiness and less sensitive to quantity compared with numeric estimates, supporting our proposed process. It is worth noting that differences in quantity still inform magnitude estimates; that is, magnitude estimates are not insensitive to quantity information, but they are more sensitive to type than they are to quantity. This experiment also demonstrates that when type and quantity do not diverge, magnitude and numeric estimates do move together—people estimated fewer calories in a healthier, smaller food portion than a less healthy, larger food portion regardless of estimation mode.

Our final two experiments provide additional process evidence via moderation, reverting to the paradigm used in experiments 1-3 with participants forming magnitude (vs. numeric estimates) of stimuli that trade-off on type versus quantity (e.g., a healthier, larger food portion vs. a less healthy, smaller food portion).

EXPERIMENT 5: INCREASING THE PRIMACY OF QUANTITY MODERATES THE EFFECT OF TYPE/QUANTITY ON MAGNITUDE ESTIMATES

Experiment 5 tested for moderation by making quantity information primary. If those forming magnitude estimates spontaneously anchor on type and fail to adjust much based on quantity (whereas those forming numeric estimates already incorporate quantity to a

considerable extent), then the divergence between estimation mode and type/quantity should attenuate when quantity information is made more primary (hypothesis 2). Using stimuli from experiment 1, we added a manipulation of whether portion size information was made primary or not. We predicted a three-way interaction between estimation mode, type/quantity, and quantity primacy. In the control condition (in which we did not make quantity information primary), we expected to observe our key interaction effect; people would estimate fewer calories in a healthier, larger food portion (vs. a less healthy, smaller food portion) when forming magnitude estimates, which would reverse when forming numeric estimates. However, we expected this effect to attenuate when increasing the primacy of quantity information, such that people would estimate more calories in a healthier, larger food portion (vs. less healthy, smaller food portion) regardless of estimation mode (hypothesis 2).

Method

We pre-registered this experiment and recruited 801 participants from Prolific ($n = 200$ per between-subjects cell). As pre-registered, we excluded those with numeric estimates 3 SD above the mean, leaving 788 ($M_{\text{age}} = 33.98$, $SD = 12.95$; 52.3% female; $M_{\text{BMI}} = 26.65$).

Participants were randomly assigned to condition in a 2 (estimation mode: magnitude vs. numeric estimate; between-subjects) \times 2 (control vs. quantity primacy; between-subjects) \times 2 (type/quantity: healthier, larger vs. less healthy, smaller; within-subject) mixed model design.

To increase the primacy of quantity information, participants in the quantity-primacy condition first rated the portion size of 12 different snack plates on a slider scale anchored at “very small” to “very large.” We presented two portion sizes (smaller vs. larger) for pretzels, chocolate covered pretzels, granola bars, chocolate covered granola bars, raisins, and chocolate covered raisins in randomized order (see examples in the web appendix).

All participants then viewed a photo of a snack plate with a larger portion of almonds (33g portion size; 200 calories) and a snack plate with a smaller portion of chocolate-covered almonds (20g portion size; 100 calories); order counterbalanced. In the magnitude estimate condition, we asked “How many calories do you think this snack plate has?” on a 9-point scale anchored at “very few calories” to “very many calories.” In the numeric estimate condition, we asked the same question as an open numeric response. We included perceived healthiness and quantity manipulation checks as in previous experiments.

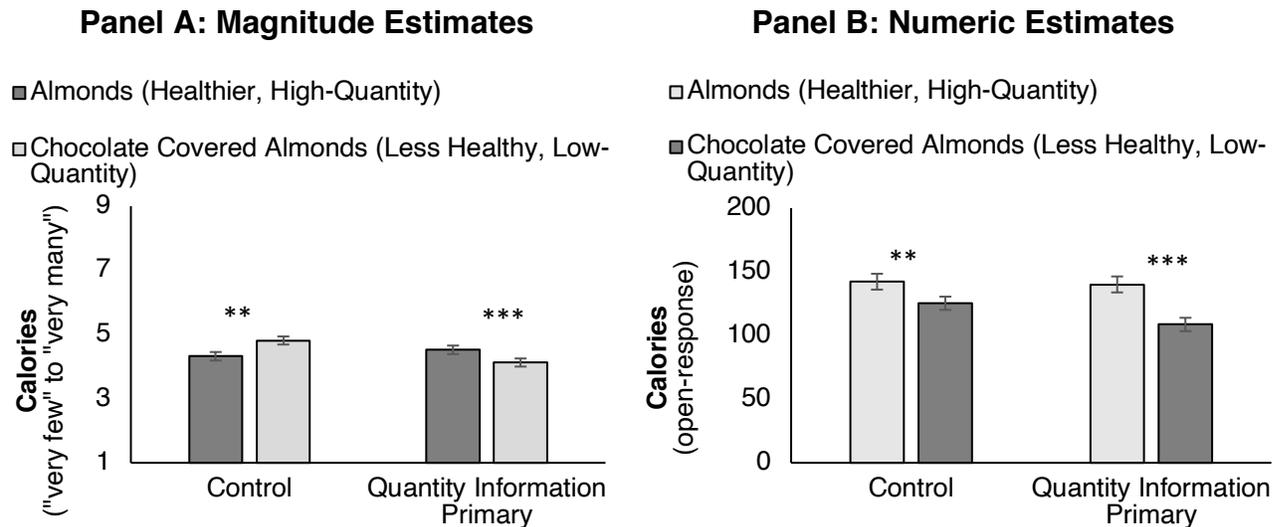
Results

Manipulation checks. We first confirmed that participants perceived the larger portion of almonds (vs. smaller portion of chocolate covered almonds) as healthier ($M_{\text{almonds}} = 7.53$; $M_{\text{chocolate covered almonds}} = 4.75$; $t(787) = 39.72, p < .001, d = 1.41$) and larger ($M_{\text{almonds}} = 5.56$; $M_{\text{chocolate covered almonds}} = 2.66$; $t(787) = 49.14, p < .001, d = 1.75$).

Magnitude estimates. We conducted a repeated measures ANOVA on calorie estimates with quantity primacy as a between-subjects factor and magnitude estimates for type/quantity as a within-subject factor, which resulted in a significant interaction ($F(1, 394) = 18.34, p < .001, \eta_p^2 = .04$; figure 4A). In line with our previous experiments, participants forming magnitude estimates estimated significantly fewer calories in a healthier, larger food portion (vs. less healthy, smaller food portion) ($M_{\text{almonds}} = 4.32, SD = 1.99$; $M_{\text{chocolate covered almonds}} = 4.81, SD = 1.87$; $F(1, 394) = 11.33, p < .001, \eta_p^2 = .03$; non-parametric Wilcoxon test: $z = 2.84, p = .004$). However, when quantity information was made primary, this effect reversed ($M_{\text{almonds}} = 4.52, SD = 1.75$; $M_{\text{chocolate covered almonds}} = 4.13, SD = 1.69$; $F(1, 394) = 7.25, p = .007, \eta_p^2 = .02$; non-parametric Wilcoxon test: $z = -2.77, p = .006$).

Numeric estimates. As predicted, opposite the magnitude estimate results, numeric estimates were significantly greater for almonds (vs. chocolate covered almonds) ($M_{\text{almonds}} = 141.15$, $SD = 87.38$; $M_{\text{chocolate covered almonds}} = 117.03$, $SD = 73.39$; $F(1, 390) = 34.76$, $p < .001$, $\eta_p^2 = .08$; non-parametric Wilcoxon test: $z = -5.73$, $p < .001$). There was a marginal interaction ($F(1, 390) = 3.13$, $p = .078$, $\eta_p^2 < .01$; figure 4B), such that this pattern was stronger in the quantity-primacy condition ($F(1, 390) = 29.08$, $p < .001$, $\eta_p^2 = .07$; non-parametric Wilcoxon test: $z = -5.42$, $p < .001$) than in the control condition ($F(1, 390) = 8.60$, $p = .004$, $\eta_p^2 = .02$; non-parametric Wilcoxon test: $z = -2.74$, $p = .006$), with no significant main effect of quantity-primacy condition ($F(1, 390) = 1.79$, $p = .182$, $\eta_p^2 < .01$).

FIGURE 4: MAGNITUDE (VS. NUMERIC) ESTIMATES ARE MORE AFFECTED BY A MANIPULATION THAT MAKES QUANTITY INFORMATION PRIMARY (EXP 5)

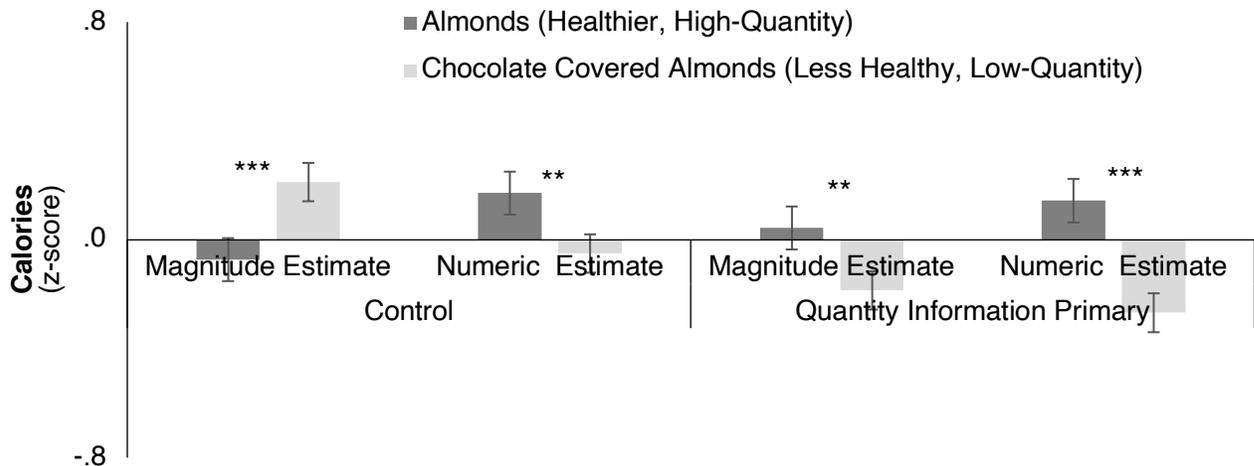


Note. ** $p < .01$; *** $p < .001$. Bars indicate ± 1 SEM.

Interaction between estimation mode and type/quantity. A repeated measures ANOVA of standardized calorie estimates on estimation mode \times quantity primacy \times type/quantity yielded a significant three-way interaction ($F(1, 784) = 4.06$, $p = .044$, $\eta_p^2 = .005$; figure 5). This

interaction demonstrates that the effect of type/quantity on magnitude estimates reversed as a function of quantity primacy, with no corresponding effect for numeric estimates.

FIGURE 5: DIVERGENCE IN STANDARDIZED MAGNITUDE AND NUMERIC ESTIMATES ATTENUATES WHEN QUANTITY INFORMATION IS MADE PRIMARY (EXPERIMENT 5)



Note. ** $p < .01$; *** $p < .001$. Bars indicate ± 1 SEM.

Discussion

Experiment 5 provided process evidence by replicating and moderating our key interaction effect. In a control condition mapping onto experiments 1-3, we replicated the divergence in magnitude versus numeric estimates. However, when quantity information was made primary, magnitude estimates mirrored numeric estimates. This experiment thus supports our proposed process, showing that those forming magnitude (vs. numeric) estimates anchor on type and do not adjust as much for quantity, unless the primacy of quantity information is increased (hypothesis 2).

EXPERIMENT 6: INTUITIVE (VS. DELIBERATIVE) MINDSET MODERATES THE EFFECT OF FOOD TYPE/QUANTITY ON NUMERIC ESTIMATES

Whereas experiment 5 identified a moderator that affects magnitude estimates, experiment 6 tested a moderator that we anticipated would affect numeric estimates. Drawing on scale compatibility, we have proposed that magnitude estimates rely primarily on type, which has primacy of processing and is more compatible with the magnitude estimate response scale (e.g., “very few” to “very many” calories). Numeric estimates instead are less compatible with type, and thus rely on continuous processing that incorporates quantity in addition to type—which necessitates careful deliberation. Experiment 6 thus tested our prediction that disrupting this deliberative mindset, by having people make numeric estimates in a quick intuitive mindset (vs. a deliberative mindset), would lead to a reversal of the pattern identified thus far for numeric estimates by requiring such estimates to rely primarily on type.

We also test whether an intuitive (vs. deliberative) mindset affects those forming magnitude estimates. In contrast to numeric estimates, we reason that magnitude estimates may be less influenced by a mindset manipulation. Food type (more vs. less healthy) is already compatible with the magnitude estimate response scale, and a deliberative mindset intervention that slows down consumers’ processing (vs. an intuitive mindset) is unlikely to prompt those forming magnitude estimates to incorporate quantity. In sum, experiment 6 tests whether intuitive (vs. deliberative) processing attenuates the effect within numeric estimates, such that these estimates exhibit the same pattern as magnitude estimates (hypothesis 3).

Method

We pre-registered this experiment and recruited 1000 participants from Prolific ($n = 250$ per between-subjects cell). As pre-registered, we excluded those who failed an attention check, leaving a final sample of 942 ($M_{\text{age}} = 33.16$, $SD = 11.95$; 49.3% female; $M_{\text{BMI}} = 25.58$).

Participants were randomly assigned to condition in a 2 (estimation mode: magnitude vs. numeric estimate; between-subjects) \times 2 (intuitive vs. deliberative mindset; between-subjects) \times 2 (type/quantity: healthier, larger vs. less healthy, smaller; within-subject) mixed model design.

Participants learned they would be making calorie judgments for different snack plates. We manipulated intuitive versus deliberative mindset by encouraging participants to make a quick, snap judgment based on their gut instincts or to take their time to provide a thoughtful, reasoned response (Rubin et al. 2019; Woolley and Risen 2018). We informed participants that we would time their decisions, and included page timers to examine processing speed (see the web appendix for response time analyses).

All participants then viewed a photo of a snack plate of a larger portion of almonds (33g portion size; 200 calories) and a snack plate of a smaller portion of chocolate-covered almonds (20g portion size; 100 calories); order counterbalanced. In the magnitude estimate condition, we asked “How many calories do you think this snack plate has?” on a 9-point slider scale anchored at “very few calories” to “very many calories.” In the numeric estimate condition, we asked the same question on a slider scale anchored at “0 calories” to “500 calories.” We used a slider scale for all conditions in this experiment to enable faster responding in the intuitive condition for those providing numeric estimates.

At the end of the experiment, we included a filler set of questions in which we embedded an attention check question, “I am careful about responding to questions and will select 3 for this answer,” which we used to carry out pre-registered exclusions. We also included a mindset

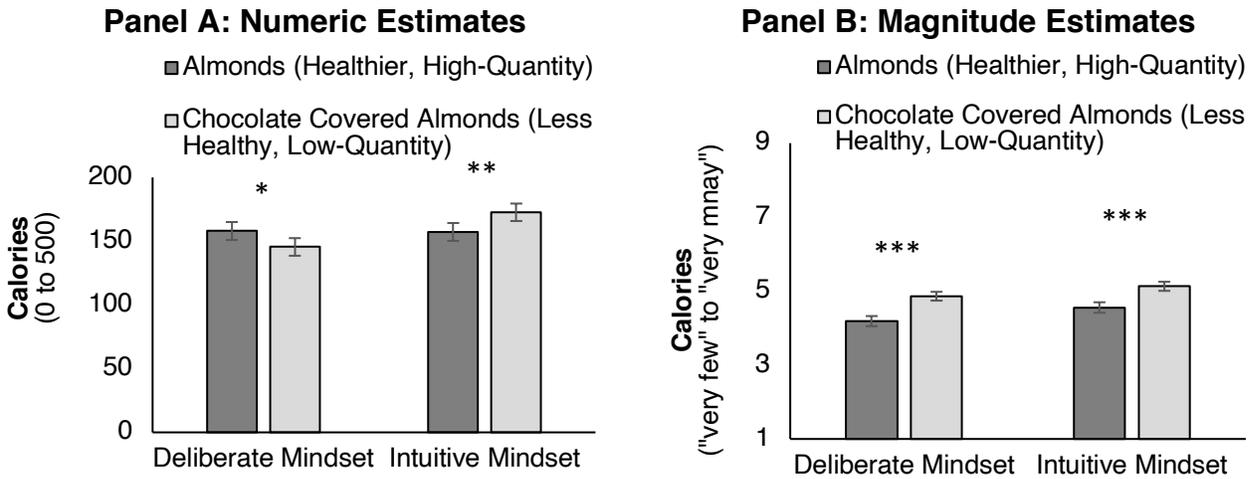
manipulation check, “What strategy did you use in indicating the number of calories in each snack plate?” with the option to select “I quickly decided based on my gut” or “I spent some time carefully deciding what answer to respond.” Most participants answered correctly (84.4%), with no significant difference by mindset condition ($\chi^2(1, N = 942) = 1.92, p = .166, \phi = .05$).

Results

Numeric estimates. We begin by examining numeric estimates, as our focal prediction is that there will be an interaction for these estimates. A repeated measures ANOVA on numeric estimates as a function of food type/quantity and mindset revealed the predicted interaction ($F(1, 472) = 11.43, p < .001, \eta_p^2 = .02$; figure 6A). Numeric estimates made in a deliberative mindset were significantly greater for almonds (vs. chocolate covered almonds) ($M_{\text{almonds}} = 157.95, SD = 104.34$; $M_{\text{chocolate covered almonds}} = 145.47, SD = 97.15$; $F(1, 472) = 4.55, p = .033, \eta_p^2 = .01$; non-parametric Wilcoxon test: $z = -1.78, p = .076$), which significantly reversed for numeric estimates made in an intuitive mindset ($M_{\text{almonds}} = 157.25, SD = 111.79$; $M_{\text{chocolate covered almonds}} = 172.64, SD = 114.09$; $F(1, 472) = 7.03, p = .008, \eta_p^2 = .02$; non-parametric Wilcoxon test: $z = 2.62, p = .009$).

Magnitude estimates. We then conducted a repeated measures ANOVA on magnitude estimates as a function of food type/quantity and mindset. As predicted, there was a significant main effect of food type/quantity such that magnitude estimates were significantly lower for almonds than for chocolate-covered almonds ($M_{\text{almonds}} = 4.39, SD = 2.10$; $M_{\text{chocolate covered almonds}} = 5.01, SD = 1.84$; $F(1, 466) = 35.15, p < .001, \eta_p^2 = .07$; non-parametric Wilcoxon test: $z = 6.02, p < .001$; figure 6B), replicating the pattern observed in prior experiments, with no significant interaction ($F(1, 466) = .18, p = .676, \eta_p^2 < .01$).

FIGURE 6: NUMERIC (VS. MAGNITUDE) ESTIMATES ARE MORE AFFECTED BY A MANIPULATION OF INTUITIVE (VS. DELIBERATIVE) MINDSET (EXPERIMENT 6)

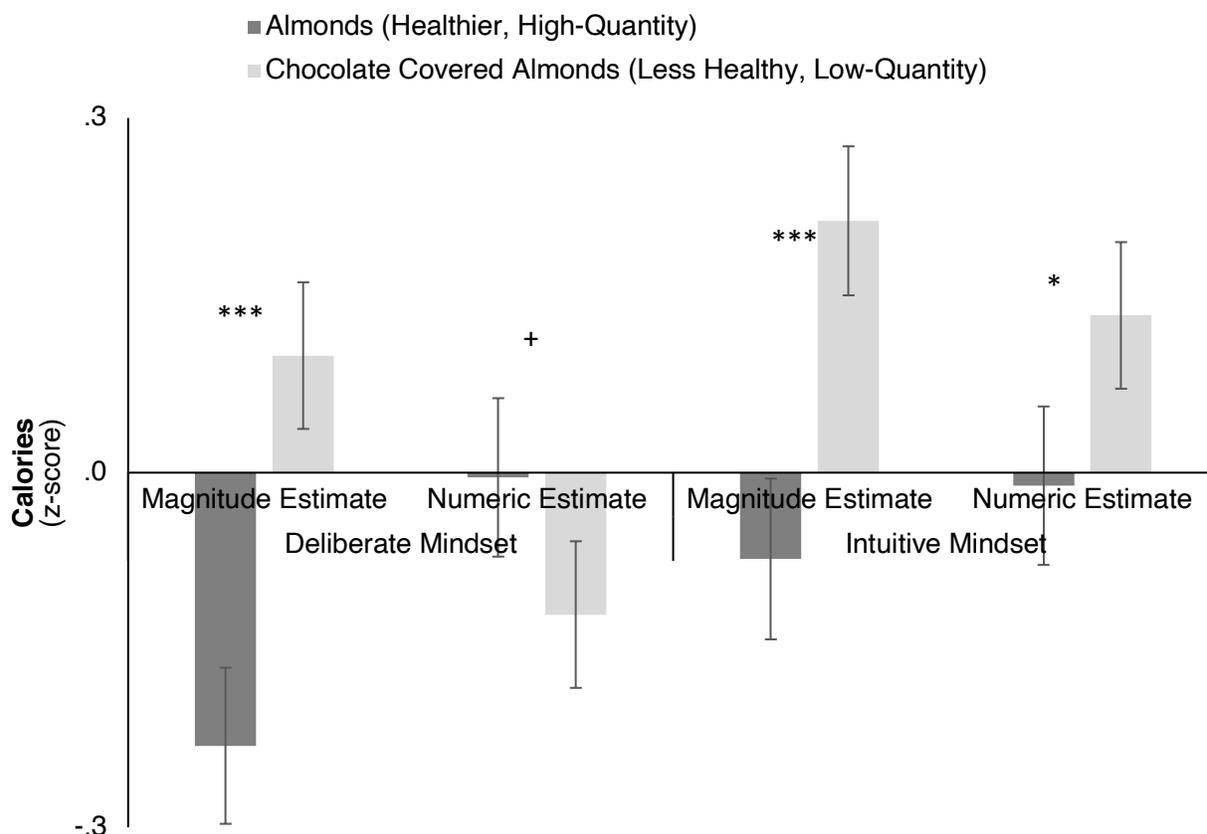


Note. * $p < .05$; ** $p < .01$; *** $p < .001$. Bars indicate ± 1 SEM.

Interaction between estimation mode and type/quantity. An analysis of the three-way interaction between estimation mode, mindset, and type/quantity was significant ($F(1, 938) = 5.53, p = .019, \eta_p^2 = .006$; figure 7), demonstrating that the effect of food type/quantity on numeric estimates significantly reverses as a function of mindset manipulation.²

² The web appendix contains a pre-registered series of three planned interaction contrasts between food type/quantity across the four estimation mode \times mindset conditions, which were all as predicted.

FIGURE 7: DIVERGENCE IN STANDARDIZED MAGNITUDE (VS. NUMERIC) ESTIMATES ATTENUATES WHEN IN AN INTUITIVE (VS. DELIBERATIVE) MINDSET (EXPERIMENT 6)



Discussion

Experiment 6 provided process evidence via moderation of numeric estimates. Consistent with our prediction that numeric estimates anchor on type (which is primary), but adjust based on quantity due to scale incompatibility, we found that having people adopt an intuitive (vs. deliberative) mindset when making numeric estimates reduced their reliance on quantity information. Indeed, in an intuitive mindset, numeric estimates closely resembled magnitude estimates. Further, we found that having people adopt a deliberative (vs. intuitive) mindset when making magnitude estimates did not affect their estimates, consistent with our proposed process

wherein magnitude estimates anchor on type (which is primary) and do not adjust much further due to scale compatibility. That is, magnitude estimates are less likely to spontaneously consider quantity information unless explicitly prompted to do so (as in experiment 5). Table 5 summarizes results across experiments 1-6.

TABLE 5: SUMMARY OF CALORIE ESTIMATES IN EXPERIMENTS 1-6.

| | | Magnitude Estimates | | Calorie Estimates | |
|------------------|------------------|------------------------------|----------------------------------|------------------------------|----------------------------------|
| | | “very few” to “very many” | | Numeric response | |
| | | Healthier, larger portion | Less healthy, smaller portion | Healthier, larger portion | Less healthy, smaller portion |
| Exp 1 (food) | | 4.03 (1.97) | 5.24 (1.70) | 116.83 (77.94) | 110.70 (59.02) |
| | | Easier, longer workout | Harder, shorter workout | Easier, longer workout | Harder, shorter workout |
| Exp 2 (exercise) | | 3.90 (1.17) | 4.88 (1.48) | 159.48 (129.86) | 129.59 (131.74) |
| | | Healthier, larger portion | Less healthy, smaller portion | Healthier, larger portion | Less healthy, smaller portion |
| Exp 3 (food) | | 5.81 (1.58) | 7.31 (1.58) | 545.37 (251.89) | 499.92 (259.71) |
| Exp 4 (food) | | Healthier | Less healthy | Healthier | Less healthy |
| | Smaller portion | 3.97 (1.95) | 5.04 (2.06) | 85.18 (55.36) | 73.04 (44.13) |
| | Larger portion | 4.79 (2.32) | 5.44 (1.64) | 100.00 (64.48) | 140.19 (87.46) |
| | | Healthier, larger portion | Less healthy, smaller portion | Healthier, larger portion | Less healthy, smaller portion |
| Exp 5 (food) | Control | 4.32 (1.99) | 4.81 (1.87) | 142.21 (90.24) | 125.28 (76.79) |
| | Quantity Primary | 4.52 (1.75) | 4.13 (1.69) | 140.07 (84.58) | 108.62 (68.94) |
| Exp 6 (food) | Deliberative | 4.24 (2.01) | 4.90 (1.89) | 157.95 (104.34) | 145.47 (97.15) |
| | Intuitive | 4.56 (2.20) | 5.13 (1.79) | 157.25 (111.79) | 172.64 (114.09) |

Note. Magnitude estimates were assessed on a nine-point Likert scale in experiments 1-5 and a nine-point slider scale in experiment 6; numeric estimates were assessed as an open numeric response for experiments 1-5 and on a slider scale from 0-500 for experiment 6.

GENERAL DISCUSSION

Six experiments showed that when assessing calories, how you estimate calories matters. We document divergences between magnitude and numeric estimates when evaluating stimuli that tradeoff type and quantity properties. Consumers forming magnitude estimates estimated fewer calories in healthier, larger food portions than less healthy, smaller food portions, which reversed or attenuated when forming numeric estimates. Similarly, they estimated fewer calories burned for easier, long workouts than for harder, shorter workouts, which significantly reversed

when forming numeric estimates. This effect emerged when less healthy, smaller food portions contained objectively fewer calories than healthier, larger food portions (experiments 1, 3-6) and when calories burned were equated across exercise intensity/quantity (experiment 2). We further demonstrate that this bias can affect food choices: When choosing a low-calorie snack plate, those forming magnitude estimates were more likely to choose a healthier, larger food portion (vs. less healthy, smaller portion) than those forming numeric estimates, even though the larger, healthier portion actually contained more calories (experiment 1).

These results also have methodological implications. Whereas researchers may measure magnitude or numeric estimates interchangeably (see table A1 in the appendix), anticipating that they move in the same direction, this is not always the case; the method of assessing calories can serve as a moderator of existing research. We replicated a prior finding from the literature when using numeric estimates, as in the initial study (Chandon and Wansink 2007), but found that the pattern reversed when using magnitude estimates (experiment 3).

This effect occurs because estimation mode influences consumers' sensitivity to information about stimuli type and quantity. Providing evidence for both components, we find that when type and quantity are fully crossed, magnitude estimates were more strongly influenced by differences in stimuli type (i.e., more vs. less healthy food) and less influenced by differences in stimuli quantity, whereas numeric estimates were less influenced by type than quantity (experiment 4). In line with these findings and adding further process insight, moderators affecting whether participants attended more to quantity (experiment 5) or type (experiment 6) eliminated this divergence. First, when participants focused on quantity information before forming their calorie estimates, magnitude estimates mirrored numeric estimates; people estimated more calories in a larger portion of healthier food than in a smaller

portion of less healthy food (experiment 5). Second, when participants made quick, intuitive judgments that focused them on what is primary about a stimulus (i.e., type) rather than slower, deliberative judgments, numeric estimates mirrored magnitude estimates; people estimated fewer calories in a larger portion of healthier food than in a smaller portion of less healthy food (experiment 6).

We note that we predicted a significant interaction between estimation mode and stimuli trading off type/quantity across all experiments, such that type/quantity would significantly affect magnitude estimates (hypothesis 1a), but that this pattern would attenuate or reverse for numeric estimates (hypothesis 1b). Indeed, whereas we observe significant effects when comparing stimuli trading off type/quantity for magnitude estimates across all experiments, in one case (experiment 1 in the lab with undergraduates), this comparison was not significant for numeric estimates. At first glance, this may seem surprising, given that we do find significant effects when assessing numeric estimates using identical stimuli from experiment 1 in experiments 4, 5, and 6. We offer the following speculation for this difference in results. Note that numeric estimates are by their nature, considerably noisier than magnitude estimates, both because the scale they are assessed on has a wider range and because we theorize that numeric estimates weigh both quantity and type (whereas magnitude estimates weigh primarily type). As a result, multiple factors may influence numeric estimates—including consumer individual differences. Indeed, experiment 1 was the only experiment conducted with a younger (i.e., undergraduate) sample, whereas experiments 4-6 all involved online samples, which skew older and have a wider age range. To the extent that older (e.g., middle-aged) participants have greater nutritional knowledge (perhaps due to greater weight-related health concerns), which improves the accuracy of numeric estimates, differences in age (and associated characteristics) may

explain these different findings. Altogether, however, we are able to conclude that whether consumers use magnitude versus numeric estimates matters across experiments. In the food domain, those forming magnitude estimates estimated smaller, unhealthier portions as having more calories than larger, healthier portions; this pattern then attenuates (experiment 1) and even reverses (all other experiments) for numeric estimates.

Theoretical Implications

This work offers theoretical insights for understanding how two key informational inputs into calorie estimates (type and quantity) are differentially weighted as a function of estimation mode. In doing so, our research adds to the literature on preference reversals (Tversky et al. 1988; Tversky, Slovic, and Kahneman 1990), which has shown that different preference elicitation modes (e.g., choice vs. WTP) can lead to different outcomes by causing people to anchor on different types of information (Moon and Nelson 2019; O'Donnell and Evers 2019; Shennib, Catapano, and Levav 2019). Analogously, we identify a divergence in calorie estimates as a function of estimation mode by causing people to anchor more on type information or to adjust from type to consider quantity to a greater extent as a function of calorie estimation mode.

As the preference reversal literature often implicates a scale compatibility process, our work also offers a new bridge between the considerable literature on calorie judgments (Campbell and Warren 2015; Chernev and Chandon 2011; Liu et al. 2019) and literature on scale compatibility. Specifically, scale compatibility suggests that people rely on information most compatible with the units of the response scale (Tversky et al. 1988). In our food experiments, although information on perceived healthiness (i.e., type) was not explicitly presented to participants on a magnitude estimate scale, we suggest that people spontaneously reason about

type categorically (i.e., “easier” to “harder” workout; “less healthy” to “very healthy” food), which is compatible with the magnitude estimate response scale, but not with the numeric estimate response scale. The lack of scale compatibility when forming numeric estimates thus prompts incorporating quantity information and requires a deliberative mindset.

Relatedly, our research also connects to the literature on scaling effects. For example, research has demonstrated systematic differences in eliciting responses on an open response scale versus a sliding scale, such that those using the latter response format bid more (Thomas and Kyung 2019). Further, boundedness can influence estimations of quantity, such that numerical estimates are more accurate for decreasing (vs. increasing) quantities, because quantity decreases are capped by a lower bound (Chandon and Ordabayeva 2017). Importantly, we demonstrate that our effects are not driven by the mechanics of the elicitation scale format (i.e., Likert vs. text box), as we find similar results when assessing magnitude estimates versus numeric estimates on bounded slider scales (experiment 6, deliberate mindset condition). Instead, our findings are driven by differences as a function of estimation mode.

Our research also offers important contributions to consumer research on calorie estimation. As table A1 in the appendix illustrates, researchers frequently use magnitude estimates and numeric estimates to assess calories, and often use them interchangeably. Thus, identifying that the mode of estimating calories matters, and understanding the cognitive processes that underlie differences in calorie estimation modes, offers a key contribution. We suggest that both estimation modes can provide insight into consumers’ calorie estimates, as consumers use both modes to assess calories (pilot study in the web appendix) and might be prompted by certain situations to use one estimation mode more than the other. Thus, conceptually analogous to calls by other researchers examining choice versus WTP (Frederick

and Loewenstein 2008; O'Donnell and Evers 2019), we call for researchers assessing calories to use both magnitude estimates and numeric estimates in measuring consumers' judgments—especially in situations in which type and quantity trade-off.

Marketing and Policy Implications

Our findings suggest that when consumers estimate calories in the food they eat, the estimation mode they use has consequences not just for their calorie assessment, but also for the amount and type of food they choose for a snack. Consumers should thus be aware that when they form magnitude estimates rather than numeric estimates, they may overweight food type relative to quantity. For example, when choosing what to eat for lunch, they may attend to perceived food healthiness (salad vs. burger) and neglect portion size. This has implications for policymakers who are actively working on tools to provide nutrition information and guidelines to consumers. Simply recommending that people monitor their caloric intake and expenditure might lead to different outcomes, depending on which estimation mode consumers use.

This research is also relevant to marketers of weight management programs or technologies that track calories to help consumers with their weight or health. Marketers for programs that facilitate numeric estimates might emphasize the need to consider the relative healthiness of the food item or intensity of the exercise workout in circumstances in which type might play a considerable role in improving health (e.g., when doing some high-intensity exercise has cardiovascular benefits). Conversely, marketers for programs or health metrics that focus on magnitude estimates might emphasize to consumers that simply forming an estimate from “very few” to “very many” calories may lead to an underweighting of quantity inputs.

Finally, if marketers or policymakers want to encourage calorie-conscious consumers to prioritize quantity (vs. type) considerations, our research suggests that they ought to encourage a focus on numeric estimates (vs. magnitude estimates). At times, having a smaller portion of a less healthy food (e.g., a small piece of chocolate) might be better from a calorie consumption perspective than having a large portion of perceptually healthier food (e.g., a large bag of pretzels). If aiming to encourage increased acceptance of smaller portions of tasty indulgences (Cornil and Chandon 2016; Liu et al. 2015), marketers might prompt calorie-conscious consumers to focus on numeric estimates.

Avenues for Future Research

This research is the first to propose and examine a difference between two common modes for assessing calories and opens the door to additional research questions. We identify three broad sets of future research questions here. First, whereas we focused primarily on calorie assessments, future research could explore whether our theory generalizes to assessments of other metrics besides calories (e.g., fat, protein, sugar, or carbohydrate intake for a given food portion). Moreover, beyond attributes related to food (or exercise), an exciting avenue for future investigation is examining whether other attributes involving both type and quantity as informational inputs may be similarly affected as a function of which estimation mode is being used. We posit that the underlying psychology for how type versus quantity information affects judgments is likely to apply more broadly outside of calorie assessments. For example, when judging the price of something, does forming a magnitude estimate (“less expensive” to “more expensive”) diverge from forming a numeric estimate (numerical response of amount spent) as a function of type (i.e., brand tier or perceived quality) and size/amount?

Second, future research is needed to more fully examine the consumer welfare implications of this discrepancy between magnitude estimates and numeric estimates. Although the current research did not assess which estimation mode is objectively better, either in terms of “accuracy” or for overall consumer health, we note that in some situations, forming magnitude estimates may lead consumers to choose higher calorie food portions (experiment 1). Further, our experiments suggest that numeric estimates may better reflect objective differences between food type/quantity in terms of calorie amounts, at least when such judgments are made with some deliberation. Indeed, epidemiologists in nutrition commonly use numeric estimates, as this estimation mode fits the goal of estimating total calorie intake and expenditure.

Future research should further examine how formation of magnitude versus numeric estimates compares to other methods of evaluating the healthiness of the food we eat in terms of which produces better outcomes for overall consumer health (Wilson 2019). Nutritionists and policy makers currently offer conflicting advice, either to focus on numeric calorie content or consider whether a given food portion is “low” or “high” in calories (Langer 2017; Wilson 2019). For example, the U.S. uses calorie labeling as a policy tool to curb obesity (CDC 2015). Yet in France and other EU countries, the Nutri-Score labeling system provides a single summary indicator on a scale from A to E, colored from green to dark orange, corresponding to the amount of positive and negative nutrients in the food item (Dubois et al. 2020; Julia and Hercberg 2017). It would also be informative to understand whether people are motivated to use one estimation mode over another to license indulgence. For example, to indulge in a small brownie, perhaps people may focus on a numeric estimate (i.e., this snack contains 100 calories) rather than a magnitude estimate (i.e., this snack contains many calories), a possibility which future research can test.

Lastly, marketing messages, product packaging, and labeling can influence the attention people pay to food type and quantity. Marketers can emphasize perceived healthiness of food through front-of-package claims (André, Chandon, and Haws 2019) or can emphasize amount of food by using partitions to make portion size more salient (Cheema and Soman 2008; Geier, Wansink, and Rozin 2012). Likewise, policymakers can emphasize perceived healthiness of food through front-of-package labels such as Nutri-Score (Dubois et al. 2020) or amount of food through interventions like “MyPlate” (Ratner and Riis 2014). One possibility is that messages emphasizing type (vs. quantity) affect magnitude (vs. numeric) estimates more, such that consumers trying to lose or maintain weight by forming numeric estimates may be less affected by such messages. Another possibility is that the use of these messages could in themselves serve as antecedents that affect whether consumers use one estimation assessment mode or the other—which, as we show, matters.

APPENDIX

TABLE A1: EXAMPLES OF RESEARCH MEASURING THE CALORIES CONSTRUCT

| <i>Research using one measure of calories – either magnitude or numeric estimates</i> | |
|---|--|
| Research assessing calories with a magnitude estimate | Research assessing calories with a numeric estimate |
| Charbonnier et al. (2016): Participants viewed images of high and low calorie food portions and answered “How many calories do you think the product consists of?” (1 = “very few calories”; 9 = “many calories”). | Biswas et al. (2014): Participants provided a numerical estimate of the caloric content of the meal. |
| Dohle, Rall, and Siegrist (2016): Participants received a recipe for a milkshake, received the milkshake itself, and tasted the milkshake. To measure participants’ estimates of the milkshake’s energy content, they were asked to judge the calories in the shake (1 = “no calories”; 9 = “many calories”). | Campbell and Warren (2015): “How many calories do you think are in this scoop of ice cream?” and “How many calories do you think you still need to burn in order to reach your goal?” (<i>numeric response</i>). |
| Geiselman et al. (1998): Participants viewed and tasted four puddings and answered “How many calories you think this food has” using a 100-mm scale (“very few calories”; “very many calories”). | Carels, Harper, and Konrad (2006): “What is your estimate of the calories in the serving above?” (<i>numeric response</i>). |
| Labiner-Wolfe, Lin, and Verill (2010): Participants viewed an image of a product (loaf of bread; single-serving frozen beef dinner with vegetables) and answered “How high or low do you consider this product to be in calories?” (“very low”; “very high”). | Chandon and Ordabayeva (2017): Participants provided a numerical estimate for the number of calories contained in different-sized portions of foods (supplemental studies 2-3). |
| Manippa et al. (2017): Participants viewed images of high and low calorie food portions from Charbonnier et al. 2016 and answered “How many calories do you think this product consists of?” (1 = “very few calories”; 9 = “many calories”). | Chandon and Wansink (2007): Consumers who had just finished eating at McDonald’s or Subway provided a numerical estimate of the caloric content of their meal. |
| Morales and Fitzsimons (2007): Participants observed a rice cake product and rated the number of calories (1 = “very few calories”; 10 = “very high in calories”). | Chernev (2011): Participants provided a numerical estimate of the caloric content of the meals. |
| Schuldt and Schwarz (2010): “Compared to other cookie brands, do you think that 1 serving of these [organic] Oreo cookies contains fewer calories or more calories?” (1 = “fewer calories”; 7 = “more calories”). | Chernev and Gal (2010): Participants provided a numerical estimate of the caloric content of the meal. |
| Slotterback, Leeman, and Oakes (2006) “Participants were instructed to rate each physical activity in terms of the total amounts of calories used.” For example: “two hours of gardening would use” (1 = “very few calories”, 4 = “moderate amount of calories”, 7 = “many calories”). | Cornil et al. (2014): Participants provided a numerical estimate of the caloric content in various food portions. |
| | Horne et al. (2019): “Estimate the total number of calories in the image using only one of two slides” Visual analog scales ranging from (0 to 999, 1000-11000 calories). |
| | Jiang and Lei (2014): Participants viewed a description and picture of a food and provided a numerical estimate of the caloric content. |
| | Salerno and Sevilla (2019): Participants estimated calories in a kiwi with a slider ranging from 1 to 400. |
| <i>Research using both measures of calories – both magnitude and numeric estimates</i> | |
| Dohle, Rall, and Siegrist (2014): “Participants were asked to judge a milkshake” (1 = “no calories”; 9 = “many calories”). | Dohle et al. (2014): Participants provided a numeric estimate of the food. |
| Liu et al. (2019): “How many calories do you think this snack has?” (1 = “very few calories”; 9 = “a lot of calories”). | Liu et al. (2019): “How many calories do you estimate are in this snack?” (<i>numeric free response</i>). |

Note: References for Appendix Table A1 are in the web appendix.

REFERENCES

- André, Quentin, Pierre Chandon, and Kelly Haws (2019), "Healthy through Presence or Absence, Nature or Science?: A Framework for Understanding Front-of-Package Food Claims," *Journal of Public Policy & Marketing*, 38 (2), 172-91.
- Borresen, Kelsey (2020), "Counting Calories Is Basically Pointless. So Why Are We Still Doing It?," https://www.huffpost.com/entry/counting-calories-pointless_1_5e18c0e2c5b6da971d1579dd.
- Campbell, Margaret C. and Caleb Warren (2015), "The Progress Bias in Goal Pursuit: When One Step Forward Seems Larger Than One Step Back," *Journal of Consumer Research*, 41 (5), 1316-31.
- Carels, Robert A., Jessica Harper, and Krista Konrad (2006), "Qualitative Perceptions and Caloric Estimations of Healthy and Unhealthy Foods by Behavioral Weight Loss Participants," *Appetite*, 46 (2), 199-206.
- CDC (2015), "Adult Obesity Facts," <https://www.cdc.gov/obesity/data/adult.html>.
- Chaiken, Shelly and Yaacov Trope (1999), *Dual-Process Theories in Social Psychology*, New York: Guilford Press.
- Chandon, Pierre and Brian Wansink (2007), "The Biasing Health Halos of Fast-Food Restaurant Health Claims: Lower Calorie Estimates and Higher Side-Dish Consumption Intentions," *Journal of Consumer Research*, 34 (3), 301-14.
- Charbonnier, Lisette, Floor van Meer, Laura N. van der Laan, Max A. Viergever, and Paul A. M. Smeets (2016), "Standardized Food Images: A Photographing Protocol and Image Database," *Appetite*, 96, 166-73.
- Cheema, Amar and Dilip Soman (2008), "The Effect of Partitions on Controlling Consumption,"

- Journal of Marketing Research*, 45 (6), 665-75.
- Chernev, Alexander (2011), "The Dieter's Paradox," *Journal of Consumer Psychology*, 21 (2), 178-83.
- Chernev, Alexander and Pierre Chandon (2011), "Calorie Estimation Biases in Consumer Choice," in *Leveraging Consumer Psychology for Effective Health Communication*, ed. Rajeev Batra, Punam A. Keller and Victor Strecher, Armonk, NY: M.E. Sharpe, 104-21.
- Chernev, Alexander and David Gal (2010), "Categorization Effects in Value Judgments: Averaging Bias in Evaluating Combinations of Vices and Virtues," *Journal of Marketing Research*, 47 (4), 738-47.
- Cornil, Yann and Pierre Chandon (2016), "Pleasure as a Substitute for Size: How Multisensory Imagery Can Make People Happier with Smaller Food Portions," *Journal of Marketing Research*, 53 (5), 847-64.
- Dohle, Simone, Sina Rall, and Michael Siegrist (2014), "I Cooked It Myself: Preparing Food Increases Liking and Consumption," *Food Quality and Preference*, 33, 14-16.
- Dubois, Pierre, Paulo Albuquerque, Olivier Allais, Céline Bonnet, Patrice Bertail, Pierre Combris, Saadi Lahlou, Natalie Rigal, Bernard Ruffieux, and Pierre Chandon (2020), "Effects of Front-of-Pack Labels on the Nutritional Quality of Supermarket Food Purchases: Evidence from a Large-Scale Randomized Controlled Trial," *Journal of the Academy of Marketing Science*, forthcoming.
- FDA (2004), "Calories Count: Report of the Working Group on Obesity,"
<https://wayback.archive-it.org/7993/20170406030602/https://www.fda.gov/Food/FoodScienceResearch/ConsumerBehaviorResearch/ucm081696.htm>.

- Fischer, Gregory W., Ziv Carmon, Dan Ariely, and Gal Zauberman (1999), "Goal-Based Construction of Preferences: Task Goals and the Prominence Effect," *Management Science*, 45 (8), 1057-75.
- Fischer, Gregory W. and Scott A. Hawkins (1993), "Strategy Compatibility, Scale Compatibility, and the Prominence Effect," *Journal of Experimental Psychology: Human Perception and Performance*, 19 (3), 580-97.
- Frederick, Shane and George Loewenstein (2008), "Conflicting Motives in Evaluations of Sequences," *Journal of Risk and Uncertainty*, 37 (2-3), 221-35.
- Geier, Andrew, Brian Wansink, and Paul Rozin (2012), "Red Potato Chips: Segmentation Cues Can Substantially Decrease Food Intake," *Health Psychology*, 31 (3), 398-401.
- Hartmann-Boyce, Jamie, David J. Johns, Susan A. Jebb, Paul Aveyard, and Behavioural Weight Management Review Group (2014), "Effect of Behavioural Techniques and Delivery Mode on Effectiveness of Weight Management: Systematic Review, Meta-Analysis and Meta-Regression," *Obesity Reviews*, 15 (7), 598-609.
- Hsee, Christopher K. (1996), "The Evaluability Hypothesis: An Explanation for Preference Reversals between Joint and Separate Evaluations of Alternatives," *Organizational Behavior and Human Decision Processes*, 67 (3), 247-57.
- Julia, Chantal and Serge Hercberg (2017), "Development of a New Front-of-Pack Nutrition Label in France: The Five-Colour Nutri-Score," *Public Health Panorama*, 3 (04), 712-25.
- Kahneman, Daniel and Shane Frederick (2002), "Representativeness Revisited: Attribute

- Substitution in Intuitive Judgment," in *Heuristics of Intuitive Judgment: Extensions and Applications*, ed. Thomas Gilovich, Dale Griffin and Daniel Kahneman, New York: Cambridge University Press.
- Langer, Abby (2017), "Here's Why Counting Calories Really Isn't Necessary for Weight Loss," <https://www.self.com/story/counting-calories-not-necessary-weight-loss>.
- Liu, Peggy J., Kelly L. Haws, Cait Lambertson, Troy H. Campbell, and Gavan J. Fitzsimons (2015), "Vice-Virtue Bundles," *Management Science*, 61 (1), 204-28.
- Liu, Peggy J., Kelly L. Haws, Karen Scherr, Joseph P. Redden, James R. Bettman, and Gavan J. Fitzsimons (2019), "The Primacy of 'What' over 'How Much': How Type and Quantity Shape Healthiness Perceptions of Food Portions," *Management Science*, 65 (7), 3353-81.
- Moon, Alice and Leif D. Nelson (2019), "The Uncertain Value of Uncertainty: When Consumers Are Unwilling to Pay for What They Like," *Management Science*, forthcoming.
- O'Donnell, Michael and Ellen R. K. Evers (2019), "Preference Reversals in Willingness to Pay and Choice," *Journal of Consumer Research*, 45 (6), 1315-30.
- Oakes, Michael E. (2005), "Stereotypical Thinking About Foods and Perceived Capacity to Promote Weight Gain," *Appetite*, 44 (3), 317-24.
- Oakes, Michael E. and Carole S. Slotterback (2005), "Too Good to Be True: Dose Insensitivity and Stereotypical Thinking of Foods' Capacity to Promote Weight Gain," *Food Quality and Preference*, 16 (8), 675-81.
- Ordabayeva, Nailya and Pierre Chandon (2016), "In the Eye of the Beholder: Visual Biases in Package and Portion Size Perceptions," *Appetite*, 103, 450-57.
- Ratner, Rebecca K. and Jason Riis (2014), "Communicating Science-Based Recommendations

- with Memorable and Actionable Guidelines," *Proceedings of the National Academy of Sciences*, 111 (Supplement 4), 13634-41.
- Rozin, Paul, Michele Ashmore, and Maureen Markwith (1996), "Lay American Conceptions of Nutrition: Dose Insensitivity, Categorical Thinking, Contagion, and the Monotonic Mind," *Health Psychology*, 15 (6), 438-47.
- Rubin, Emily B., Anna E. Buehler, Elizabeth Cooney, Nicole B. Gabler, Adjoa A. Mante, and Scott D. Halpern (2019), "Intuitive Vs Deliberative Approaches to Making Decisions About Life Support: A Randomized Clinical Trial," *JAMA network open*, 2 (1), e187851.
- Schley, Dan R., Bart De Langhe, and Andrew R. Long (2020), "System 1 Is Not Scope Insensitive: A New, Dual-Process Account of Subjective Value," *Journal of Consumer Research*, forthcoming.
- Schwarz, Norbert, Hans-J Hippler, Brigitte Deutsch, and Fritz Strack (1985), "Response Scales: Effects of Category Range on Reported Behavior and Comparative Judgments," *Public Opinion Quarterly*, 49 (3), 388-95.
- Shennib, Fuad Adnan, Rhia Catapano, and Jonathan Levav (2019), "Preference Reversals between Digital and Physical Goods," in *Association for Consumer Research*, ed. Rajesh Bagchi, Lauren Block and Leonard Lee, Atlanta, GA.
- Slovic, Steven (2014), "Two Systems of Reasoning: An Update," in *Dual Process Theories of the Social Mind*, ed. Jeffrey W. Sherman, Bertram Gawronski and Yaacov Trope, New York: Guilford.
- Slotterback, Carole S., Heather Leeman, and Michael E. Oakes (2006), "No Pain, No Gain: Perceptions of Calorie Expenditures of Exercise and Daily Activities," *Current Psychology*, 25 (1), 28-41.

- Slovic, Paul, Dale Griffin, and Amos Tversky (1990), "Compatibility Effects in Judgment and Choice," in *Insights in Decision Making: A Tribute to Hillel J. Einhorn*, ed. Robin M. Hogarth, Chicago, IL: University of Chicago Press, 5-27.
- Spence, Ian (1990), "Visual Psychophysics of Simple Graphical Elements," *Journal of Experimental Psychology: Human Perception and Performance*, 16 (4), 683-92.
- Thomas, Manoj and Ellie J. Kyung (2019), "Slider Scale or Text Box: How Response Format Shapes Responses," *Journal of Consumer Research*, 45 (6), 1274-93.
- Thomas, Manoj and Vicki Morwitz (2009), "Heuristics in Numerical Cognition: Implications for Pricing," in *Handbook of Pricing Research in Marketing*, ed. Vithala R. Rao, Cheltenham, UK: Edward Elgar, 132-49.
- Tversky, Amos, Shmuel Sattath, and Paul Slovic (1988), "Contingent Weighting in Judgment and Choice," *Psychological Review*, 95 (3), 371-84.
- Tversky, Amos, Paul Slovic, and Daniel Kahneman (1990), "The Causes of Preference Reversal," *American Economic Review*, 204-17.
- Viswanathan, Madhubalan and Terry L. Childers (1996), "Processing of Numerical and Verbal Product Information," *Journal of Consumer Psychology*, 5 (4), 359-85.
- Wilson, Peter (2019), "Death of the Calorie,"
<https://www.economist.com/news/2019/03/16/death-of-the-calorie>.
- Windschitl, Paul D. and Gary L. Wells (1996), "Measuring Psychological Uncertainty: Verbal Versus Numeric Methods," *Journal of Experimental Psychology: Applied*, 2 (4), 343-64.
- Woolley, Kaitlin and Jane L. Risen (2018), "Closing Your Eyes to Follow Your Heart: Avoiding Information to Protect a Strong Intuitive Preference," *Journal of Personality and Social Psychology*, 114 (2), 230.