

Choosing the Light Meal: Real-Time Aggregation of Calorie Information Reduces Meal Calories

Journal of Marketing Research
2021, Vol. 58(5) 948-967
© American Marketing Association 2021
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/00222437211022367
journals.sagepub.com/home/mrj



Eric M. VanEpps , Andras Molnar, Julie S. Downs,
and George Loewenstein

Abstract

Numeric labeling of calories on restaurant menus has been implemented widely, but scientific studies have generally not found substantial effects on calories ordered. The present research tests the impact of a feedback format that is more targeted at how consumers select and revise their meals: real-time aggregation of calorie content to provide dynamic feedback about meal calories via a traffic light label. Because these labels intuitively signal when a meal shifts from healthy to unhealthy (via the change from green to a yellow or red light), they prompt decision makers to course-correct in real time, before they finalize their choice. Results from five preregistered experiments ($N = 11,900$) show that providing real-time traffic light feedback about the total caloric content of a meal reduces calories in orders, even compared with similar aggregated feedback in numeric format. Patterns of ordering reveal this effect to be driven by people revising high-calorie orders more frequently, leading them to choose fewer and lower-calorie items. Consumers also like traffic light aggregation, indicating greater satisfaction with their order and greater intentions to return to restaurants that use them. The authors discuss how dynamic feedback using intuitive signals could yield benefits in contexts beyond food choice.

Keywords

calorie labeling, food ordering, aggregation, traffic light, dynamic feedback, revisions, intuitive

Online supplement: <https://doi.org/10.1177/00222437211022367>

Modern consumers have easy access to more information than ever before, yet this has not led to measurably better decisions in domains, such as diet, that require self-control. One central problem is that information provided to consumers is often complicated, requiring them to integrate information in ways that exceed their abilities or motivations. That is, information may be available, but it is not always *actionable*. In this article, we study how feedback provided in an intuitive form can enhance individuals' decisions by prompting them to course-correct in real time, before they finalize their choices. In particular, we investigate intuitive, dynamic feedback that responds to consumers' preliminary decisions. Such labeling provides meaningful guidance (e.g., whether to continue with the decision as planned or stop and make changes) in response to active selections and can help consumers take self-beneficial actions.

The most effective labels are intuitive, using visual elements that leverage existing strongly held associations, such as the association between a red light and stopping or between emojis and emotional valence. Feedback should be carefully calibrated to match the representations that people find

meaningful (e.g., positive vs. negative feedback categories), rather than leaving translational work to consumers. In addition, feedback should be calibrated to the decision process itself. Although static information may operate well in some contexts (e.g., when consumers have time to fully consider every option), more dynamic feedback, such as the cumulative price of goods in a shopping cart, may be more beneficial in a wide range of situations (Sheehan and Van Ittersum 2018). We build on prior literature that conceptualizes the consumer decision process as one that often relies on constructed preferences

Eric M. VanEpps is Assistant Professor of Marketing, David Eccles School of Business, University of Utah, USA (email: eric.vanepps@eccles.utah.edu). Andras Molnar is Principal Researcher (postdoctoral), Booth School of Business, University of Chicago, USA (email: andras.molnar@chicagobooth.edu). Julie S. Downs is Associate Professor of Psychology, Department of Social and Decision Sciences, Carnegie Mellon University, USA (email: downs@andrew.cmu.edu). George Loewenstein is Herbert A. Simon Professor of Economics and Psychology, Department of Social and Decision Sciences, Carnegie Mellon University, USA (email: gl20@andrew.cmu.edu).

(Bettman, Luce, and Payne 1998) that are calculated in the moment (Payne, Bettman, and Schkade 1999). Dynamically updated information about the content of one's current selections can help consumers update their preferences. We test the role of such dynamic, intuitive feedback in the context of meal choice and calorie labeling, a domain in which information provision is ubiquitous but has failed to show consistently positive effects on consumer behavior.

Calorie Labeling

As part of the 2010 Affordable Care Act, chain restaurants with 20 or more locations are now required by federal law to display numeric calorie counts on menus (Block 2018), but this numeric calorie labeling has generally failed to substantially reduce calorie consumption in real-world settings such as fast-food restaurants (Cantor et al. 2015; Downs et al. 2013; Elbel et al. 2009, 2013; Finkelstein et al. 2011; Schwartz et al. 2012). Although studies have observed a small impact of numeric calorie menu labeling on reducing calorie consumption in some settings (Auchincloss et al. 2013; Bollinger, Leslie, and Sorenson 2011; VanEpps, Downs, and Loewenstein 2016; Wisdom et al. 2010), comprehensive reviews (Bleich et al. 2017) and meta-analyses (Cadario and Chandon 2020) have failed to identify a consistent benefit of labeling.

There are numerous possible reasons for this failure. Customers may willfully ignore or not care about calorie information (Howlett et al. 2009). Labels may backfire when calorie counts are smaller than consumers expect (Burton et al. 2015; Tangari et al. 2019). Marketing gimmicks such as “just-below” calorie labeling (e.g., 799 calories vs. 800; Choi, Li, and Samper 2019) or the use of calorie ranges (e.g., a burrito with a range of 410–1,185 calories; Liu, Bettman, et al. 2015) may encourage otherwise health-motivated consumers to consume indulgent items. Or, consumers may want to reduce calories, and may try to use labels to do so, but fail to appreciate the cumulative contribution of multiple items. Indeed, previous research has shown that consumers often underestimate the calorie content of meals (Block et al. 2013), even in the presence of item-level calorie labeling (Elbel 2011). Furthermore, accuracy in calorie estimates seems to be worse among those who order larger meals (Taksler and Elbel 2014) and among those who order high-calorie beverages (Franckle, Block, and Roberto 2016), suggesting that the aggregation of this numeric information across multiple items may be difficult for consumers. For example, consumers who choose a healthy entrée may also order more calorie-rich side items or a beverage, offsetting the calories saved on the entrée (Wisdom et al. 2010). Most importantly, existing calorie label approaches do little to encourage *revisions* once an initial decision has been made.

An ideal calorie labeling system would tackle several of these factors simultaneously: it would encourage consumers to notice and care about calorie information while also prompting and enabling revisions if initial choices were too high in calories. By contrast, if initial selections were low in calories, the ideal labeling approach would provide positive feedback,

letting consumers know they can continue with their decision without need for revision. We propose a novel system—traffic light aggregation—that offers such real-time feedback and intuitive meal-level guidance, providing consumers with a salient calorie target for their meal.

Traffic Light Labels and Dynamic Aggregation

We test the impact of aggregated feedback regarding total meal calories presented in a traffic light format and updated as consumers select or remove items from their meal to provide a green-, yellow-, or red-light summary of total calories. Real-time dynamically aggregated feedback can influence current decisions by showing consumers, instantaneously, the impact of adding or subtracting an item on total meal calories, giving them both reason and opportunity to remove or substitute items to stay under identifiable calorie thresholds. This kind of dynamic aggregation applied with numeric calorie labels has shown initial promise in a hypothetical sandwich-building task (Gustafson and Zeballos 2019) but has not yet been tested for actual consumption decisions, nor has it been tested in a meal selection task where previous choices can be altered. Furthermore, in the absence of more detailed information about consumers' choice processes (e.g., timing, stopping rules, meal revisions), it is unclear from prior research *how* the presence of aggregated information might help consumers revise meal selections, a process that may vary across different types of aggregation (e.g., numeric, traffic light).

On its own, dynamic aggregation of numeric labels might not be enough to substantially influence consumer choices. U.S. Food & Drug Association guidelines currently include requirements for printing numeric calorie labels in a readable way (Federal Register 2014), but these formats do little to make calorie information visually salient or remind consumers to consider nutritional information when deciding what to order (Goswami and Urmitsky 2019). Among alternative labeling formats, traffic light icons have shown promise in attracting attention and promoting healthier choices (Downs, Wisdom, and Loewenstein 2015; Ellison, Lusk, and Davis 2013, 2014; Sonnenberg et al. 2013; Thorndike et al. 2012, 2014; VanEpps et al. 2016) by providing clear, credible signals to help people avoid overconsumption.

Previous research, however, has applied traffic light labels only to individual items in a static format and found that these labels have largely performed as well as numeric calorie labeling, but not substantially better (VanEpps et al. 2016). This highlights important practical limitations of item-level traffic light labels. For example, consumers may mistakenly believe that they have successfully met a health goal by including a green-light item as part of their order, licensing them to add additional, potentially unhealthier, items (Fishbach and Dhar 2005; Khan and Dhar 2006; Wilcox et al. 2009). Conversely, items labeled with a red light may seem strictly prohibited, undermining the credibility of labeling for those who correctly believe that those items can be acceptable in a meal otherwise composed of low-calorie items. Consumers might, therefore, benefit from guidance regarding how to

combine an item such as a soda with other items to make a healthy meal. Fundamentally, such ordinal markers cannot be easily aggregated by consumers to characterize the overall healthiness of a meal. Menus that provide dynamic feedback with a summary in the form of an aggregated light label, however, can overcome this barrier and improve consumer decisions. Aggregated traffic light feedback (i.e., at the meal level) can set an intuitively meaningful calorie target (a green light) for consumers to pursue, which, in combination with existing numeric item labels that show exactly how different items contribute to total meal calories, can lead consumers who would otherwise order too many calories to revise their meal selections and order fewer calories.

How Consumers Might Reduce Meal Calories

One barrier to implementing an effective food-labeling policy is the relative lack of empirical studies that investigate *how* and *why* calorie labeling might influence food choice; most existing studies simply test whether labels ultimately change overall food consumption. Responding to this gap in the literature, we delve into the behaviors available to consumers and consider two distinct avenues through which they can reduce their total calorie intake: (1) reducing the total number of food items and (2) substituting lower-calorie items for higher-calorie items.¹ Consumers naturally tend to put greater emphasis on the latter, reducing intake by choosing lower-calorie rather than fewer items (Liu et al. 2019), but little is known about how these methods of calorie reduction may be influenced by informational interventions. In the present research, we measure both dimensions of behavior directly to assess how consumers respond to information of different types, including dynamically aggregated calorie information.

In addition, we test how the format of dynamically aggregated feedback about meal calories—numeric calorie information versus traffic light labels—affects its impact. In one previous study (Gustafson and Zeballos 2019), consumers simultaneously reduced calorie content and improved their accuracy in estimating sandwich calories when presented with aggregation of numeric calorie content. However, we argue that traffic light aggregation can be even more effective in reducing calorie content relative to numeric aggregation due to the traffic light's intuitive format. Whereas numeric aggregation guides sequential choices in the construction of a meal (e.g., consumers add the lower-calorie side after seeing how many calories are in their previously selected items), traffic light feedback provides prescriptive guidance that also motivates revisions of existing selections when the meal's calorie

content is too high. That is, rather than affecting only prospective decisions about what menu items to include, we propose that real-time feedback in a format that is more intuitive also promotes consideration of which already-added items to remove. Consumers can readily proceed with a meal that receives a green light, stop and take actions to reduce calories if a meal receives a red light, and use discretion when they see a yellow light. Such feedback provides a simpler-to-interpret prescriptive guide regarding not just how many calories are in the selected meal but also how the meal compares to the number of calories people should order. The traffic light format is a translated format that provides a clear decision “signpost” (Larrick, Soll, and Keeney 2015; Ungemach et al. 2018) to help consumers attend to their meal calorie content, an attribute of the choice that might otherwise be overlooked if merely presented in numeric format.

The intuitive design of traffic light labels, then, is critical to our theory that consumers will respond to real-time feedback by revising their selections *when that feedback communicates that their current selections are problematic*. Similar to the experience of driving in the presence of a green traffic light, when consumers see a green light representative of low-calorie selections, they can continue without stopping (and their meal selection should therefore be unaffected by traffic light aggregation). However, when a green light shifts to a yellow or red light, this provides an unambiguous signal that should prompt immediate action. Here, we highlight the other fundamental feature of our intervention: the instantaneous, dynamic feedback regarding meal calorie content inherently targets consumers who need guidance the most—those who originally select meals that exceed the healthy range of calories and receive a yellow or red light as a result. We predict that these consumers will be the ones most affected by traffic light aggregation, choosing to revise their meals by dropping items or substituting lower-calorie alternatives so that the light will shift back to green and give them tacit encouragement.

We propose that these two factors—intuitive guidance and dynamic feedback—drive a revision process that has typically been ignored by static labeling approaches and can yield meaningful calorie reductions. We provide novel empirical data regarding this revision process by monitoring the timing and sequence of consumer actions as participants add and remove items from their meals. Through our use of real-time aggregation paired with intuitive traffic light labels, we add to the decision-making literature by showing how consumers respond to feedback indicating that their behavior falls outside of a prescribed range of acceptability.

Overview of Experiments

We present five experiments that test the impact of adding aggregated calorie information to a menu in the form of a dynamic meal-level traffic light label. All initial sample sizes were determined a priori, and we report all data exclusions, manipulations, and measures (Simmons, Nelson, and Simonsohn 2012). All experiments were preregistered on AsPredicted.org. The corresponding preregistrations, as well

¹ We recognize that a third avenue—limiting consumption of food ordered (e.g., saving food for a later meal)—is also possible, but due to logistical constraints, the present research did not include measures of actual consumption. We rely on prior research showing that reductions in calories ordered translate to reductions in calories consumed (Schwartz et al. 2012) but acknowledge that some individual consumers might attempt to reach calorie goals through shifts in consumption rather than ordering.

as all data collected, can be publicly accessed on ResearchBox.org (<https://researchbox.org/275>). In Experiment 1, we compare actual lunch choices made with and without an aggregated light label. In Experiment 2, we compare total calories ordered across three aggregated labeling conditions (no aggregation, numeric aggregation, and traffic light aggregation) and test for the proposed mechanism of additional revisions. In Experiment 3, we demonstrate that traffic light aggregation achieves calorie reductions compared with other labeling approaches even when all labeled menu conditions provide written guidance for how many calories should be in a meal. In Experiment 4, we show that the benefits of traffic light aggregation are specifically driven by the intuitive nature, rather than the discrete categorical nature, of traffic lights. Finally, in Experiment 5, we provide further support for the importance of the intuitiveness of the feedback consumers receive, as opposed to its simple visual salience, for helping consumers order fewer calories; we show that equally salient, but unintuitive, graphic feedback fails to reduce meal calories. After demonstrating, in Experiment 1, the impact of aggregate traffic light information in a real meal-ordering context, the remaining studies are hypothetical. As in prior research that investigated the process underlying meal choices from calorie-labeled menus (Parker and Lehman 2014), hypothetical experiments enable the sample size and necessary experimental control to examine the choice and revision process. Table 1 provides an overview of our experiments.

Experiment 1: Traffic Light Aggregation for Meal Choices

In Experiment 1, we test the impact of traffic light calorie label aggregation, presenting a menu that featured numeric calorie labels for each item, with or without an aggregated traffic light calorie label (e.g., “Total calories: [yellow light graphic]”) at the top of the menu that updated in real time as participants added or removed items from their order. University participants completed a short survey in exchange for the chance to win a free lunch of their choice, selected from a computerized menu (see Figures S1–S2 in the Web Appendix); winners received these meals at lunchtime, with sandwiches prepared by on-campus dining services.

Methods

Participants. We recruited 509 participants (49% female; mean age = 21.1 years, mean body mass index [BMI] = 23.4) from public spaces on the campus of a public U.S. university to complete a short study in exchange for a 1 in 10 chance of receiving their selected lunch for free. Recruitment occurred between 9:00 A.M. and 11:45 A.M. on days of data collection.

Procedure. Participants selected a meal from a computerized menu with half-size and full-size options of four different sandwiches (ranging from 215 to 640 calories), eight side options (40–220 calories), and eight drink options (0–240 calories).

Prices and calories listed on the menu were from the actual campus restaurant. Participants were randomly assigned to one of two conditions: numeric item labels only (“Item Labels”) or numeric item labels plus traffic light aggregation at the meal level (“Traffic Light Aggregation”).

Light feedback was determined by the calorie content of the meal: a green light would appear when the meal had fewer than 600 calories, yellow when it had between 600 and 899 calories, and red when the meal was 900 calories or more. Calorie recommendations can vary across experimental studies and expert guidelines (e.g., Downs et al. 2013; VanEpps et al. 2016); the Healthier Restaurant Meal Guidelines’ nutrient standards developed by a group of experts in 2012 set a meal calorie content threshold at 700 calories (Cohen et al. 2013), whereas other studies have used thresholds of 650 calories, 750 calories, or 800 calories. We sought to place these previously recommended calorie counts within the range of our “yellow-light” category, meaning participants would generally receive positive feedback (i.e., a green light) when their meals had fewer calories and negative feedback (i.e., a red light) when meals had more calories.²

Participants could order more than one item per category, had no restriction on how long they spent on the menu page, and could scroll over the names of the sandwiches to learn the ingredients. The price of the meal was limited to \$15.³ All actions (i.e., each addition and each removal of food items) were recorded. Following the meal selection, participants completed survey items regarding their current level of hunger and their demographics, and then rolled a ten-sided digital die to learn whether they won their chosen meal.

Results

Traffic Light Aggregation significantly reduced total calories ordered ($M = 683$, 95% confidence interval [CI] = [655, 710]) compared with Item Labels ($M = 730$, 95% CI = [702, 758]; $t(507) = 2.37$, $p = .018$, Cohen’s $d = .21$; see Figure 1, Panel A). Meanwhile, the number of items *ordered* did not significantly differ between Traffic Light Aggregation ($M = 2.87$; 95% CI = [2.79, 2.95]) and Item Labels ($M = 2.87$, 95% CI = [2.78, 2.95]; $t(507) = .09$, $p = .93$, Cohen’s $d = .01$) conditions (see Figure 1, Panel B). Traffic Light Aggregation led to a marginally significant decrease in calories per item ($M = 250$, 95% CI = [238, 263]) relative to Item Labels ($M = 268$, 95% CI = [255, 281]; $t(506) = 1.91$, $p = .057$, Cohen’s $d = .17$ (Figure 1, Panel C). All results are robust to the inclusion of demographic factors (gender, age, BMI; see Tables S1–S2 in the Web Appendix).

² We demonstrate robustness to a different set of calorie thresholds in Experiments 2–5.

³ Only 15 participants (about 3% of the sample) chose a meal between \$12 and \$15, suggesting that this limit was not too restrictive and that most orders were not driven by an attempt to maximize value from the study design.

Table 1. Overview of Experiments. Sample Size and Mean (SD) of Total Calories Ordered, by Condition.

Experiment and Subject Pool										
Condition		I	2	3 ^a	4	5	S1 ^b	S2 ^b	S3 ^b	S4 ^b
Item Label	Dynamic Meal Label	University Participants	MTurk	MTurk	Prolific	Prolific	Prolific	MTurk	University Alumni	MTurk
No	No			n = 611 986 (396)				n = 938 854 (336)	n = 449 678 (231)	
Numeric	No	n = 253 730 (225)	n = 603 920 (367)	n = 608 901 (353)	n = 600 956 (378)	n = 602 970 (366)	n = 301 748 (256)	n = 472 842 (327)	n = 223 660 (231)	n = 517 911 (376)
Numeric	No ^c									n = 471 945 (408)
Numeric	Numeric (Continuous)		n = 600 875 (360)	n = 608 903 (358)	n = 600 934 (362)			n = 472 803 (326)	n = 231 649 (232)	
Numeric	Numeric (Discrete)				n = 600 928 (360)					
Numeric	Traffic Light (Discrete)	n = 256 683 (225)	n = 600 815 (300)	n = 610 858 (326)	n = 601 871 (314)	n = 602 874 (307)	n = 301 716 (239)			n = 510 816 (310)
Numeric	Traffic Light (Discrete) ^c									n = 481 859 (336)
Numeric	Traffic Light (Continuous)				n = 601 860 (314)					
Numeric	Traffic Light (Unintuitive)					n = 602 940 (326)				
Numeric	Symbol (Intuitive)					n = 602 909 (341)				
Numeric	Symbol (Unintuitive)					n = 602 926 (347)				
Traffic Light	No							n = 473 830 (327)	n = 227 655 (236)	
Traffic Light	Traffic Light (Discrete)							n = 465 802 (309)	n = 242 632 (246)	
Total conditions		2	3	4	5	5	2	5	5	4

^aAll conditions in Experiment 3, except for the No Label condition, also featured a static numeric guideline on the bottom of the screen.^bS1–S4 indicate Supplemental Experiments 1–4. These experiments are reported in detail in the Web Appendix.^cThese conditions also featured a cognitive load manipulation during the meal choice.

Notes: Cells report sample size per condition after exclusions and mean number of calories ordered, with standard deviation in parentheses.

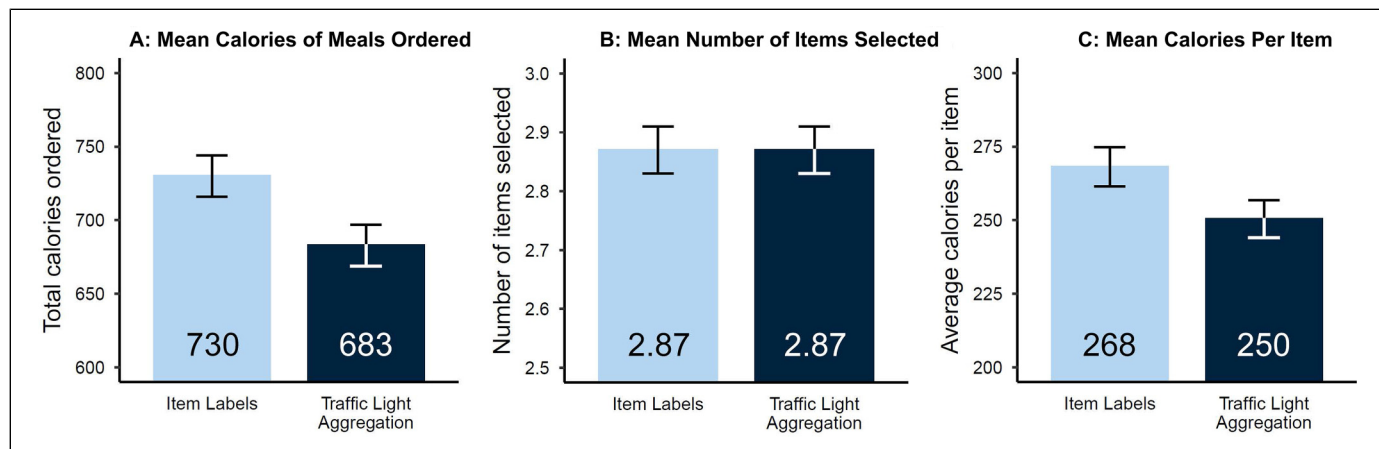


Figure 1. Experiment 1: Effect of traffic light aggregation (vs. item labels) on meals ordered.

Notes: Error bars indicate ± 1 SE.

Discussion

Examining real meal choices, Experiment 1 demonstrates a significant reduction in calories ordered by those exposed to Traffic Light Aggregation, leading to an average reduction of 47 calories per meal, or approximately 6.4% of meal calories. Experiment 1 enables us to draw conclusions about the immediate, short-term impact (i.e., calorie reduction of a single meal) of adding aggregated information to restaurant menus. To determine whether these labels could backfire in some way—for example, if consumers selected meals that they did not actually want or if they disliked being exposed to traffic light aggregation—we conducted an online follow-up study ($N = 602$, 47% female; mean age = 31.5 years) where participants ordered a hypothetical meal following the same procedures as Experiment 1 (for full details, see Web Appendix, Supplemental Experiment 1). In this study, participants reported that they would enjoy their meals significantly more, felt that the restaurant was more concerned with their well-being, and indicated that they would be more likely to return to the restaurant when the menu featured Traffic Light Aggregation as opposed to only Item Labels (all p s < .001; items adapted from Shah et al. [2014]). We speculate that the restaurant's effort to provide real-time feedback on meal healthiness demonstrates concern for consumers' health and well-being and that this positive signaling in turn leads to overall increases in customer satisfaction and intentions to return, though future research would be needed to fully test this hypothesis. Regardless, these results indicate that Traffic Light Aggregation not only achieves calorie reductions for consumers but also can help restaurants promote customer satisfaction and loyalty.

However, the more theoretical implications of Experiment 1 are somewhat limited by practical considerations. For instance, due to recruitment limitations for real-world orders, we could compare only two experimental conditions in Experiment 1. From these results, it is unclear whether traffic light aggregation is unique in its ability to produce calorie reductions, or

whether the same benefits would be achieved by an alternative approach of numeric calorie aggregation. With the larger sample size and experimental control afforded by hypothetical choice experiments, we can identify whether the combination of numeric item labels and traffic light aggregation leads to meaningfully different outcomes from other labeling approaches and can better examine the choice and revision processes that underlie these results.

Experiment 2: Comparing Traffic Light with Numeric Aggregation

In Experiment 2, we add a third condition in which numeric item labels are augmented with numeric aggregated labels. In this way, we test whether a basic type of real-time aggregation achieves sizable calorie reductions, or whether something more specific to traffic light aggregation is needed to produce the effect observed in Experiment 1. We also investigate how people respond to this meal-level feedback by adding, removing, or replacing items. Our theoretical account predicts that traffic light aggregation drives calorie reductions through an increased propensity to revise previous decisions, even relative to numeric aggregation, and so we test for mediation of any differential effect between traffic light and numeric aggregation through the number of revisions (specifically, removals of items previously added to the meal).

Methods

Participants. We recruited 1,823 participants from Amazon Mechanical Turk (MTurk) and excluded 20 participants (1%) who selected meals that cost over \$18 or contained over 2,000 calories (exclusion criteria preregistered), leaving 1,803 observations (59% female, mean age = 37.5 years, mean BMI = 27.7).

Procedure. Participants selected a hypothetical meal from a restaurant menu that featured five main dishes, five side

dishes, and five drink options (Web Appendix, Figures S3–S5). For robustness, we tested a different menu from that used in Experiment 1 and used a different set of calorie thresholds for the meal-level traffic lights (0–749 = green light, 750–1,124 = yellow light, 1,125+ = red light). Prices were marked on each item, and participants were randomly assigned to one of three calorie label conditions: labels only on items (“Item Labels”), item labels plus numeric aggregation (“Numeric Aggregation”), or item labels plus traffic light aggregation (“Traffic Light Aggregation”). Participants in all conditions were required to stay on the menu screen for at least one minute. Each item added to, and removed from, the meal was recorded in our data. Participants then completed a series of survey questions, which included demographics, estimates of the price and calorie content of their selected meal and of four preselected menu items, and subjective numeracy.

Results

Calories ordered. Numeric Aggregation ($M = 875$ calories, 95% CI = [846, 904]) and Traffic Light Aggregation ($M = 815$, 95% CI = [791, 839]) both significantly reduced total meal calories compared with Item Labels ($M = 920$, 95% CI = [890, 949]; $t(1,201) = 2.14$, $p = .032$, Cohen’s $d = .12$; $t(1,157) = 5.41$, $p < .001$, Cohen’s $d = .31$, respectively). Furthermore, Traffic Light Aggregation reduced calories compared with Numeric Aggregation ($t(1,160) = 3.12$, $p = .002$, Cohen’s $d = .18$; see Figure 2, Panel A).

Number of items. Relative to Item Labels ($M = 3.28$ items, 95% CI = [3.23, 3.34]), participants ordered fewer items when given Numeric Aggregation ($M = 3.17$, 95% CI = [3.12, 3.22]; $t(1,185) = 2.86$, $p = .004$, Cohen’s $d = .16$) or Traffic Light Aggregation ($M = 3.14$, 95% CI = [3.10, 3.18]; $t(1,143) = 3.85$, $p < .001$, Cohen’s $d = .22$), but number of items ordered did not differ between the two aggregation conditions ($t(1,183) = .92$, $p = .36$; see Figure 2, Panel B).

Average calories per item. Traffic Light Aggregation reduced per-item calories ($M = 263$ calories, 95% CI = [255, 271]) relative to both Item Labels ($M = 282$, 95% CI = [274, 290]; $t(1,196) = 3.33$, $p = .001$, Cohen’s $d = .19$) and Numeric Aggregation ($M = 277$, 95% CI = [269, 286]; $t(1,183) = 2.47$, $p = .014$, Cohen’s $d = .14$), but the latter two conditions did not differ from each other ($t(1,198) = .75$, $p = .45$; see Figure 2, Panel C).

Meal revisions: Additions and removals. Traffic Light Aggregation led to both more additions and more removals ($M_{\text{added}} = 4.17$, 95% CI = [4.02, 4.32]; $M_{\text{removed}} = 1.03$, 95% CI = [.89, 1.17]) compared with both Item Labels ($M_{\text{added}} = 3.77$, 95% CI = [3.68, 3.87]; $M_{\text{removed}} = .49$, 95% CI = [.42, .57]) and Numeric Aggregation ($M_{\text{added}} = 3.73$, 95% CI = [3.63, 3.84]; $M_{\text{removed}} = .56$, 95% CI = [.47, .65]) (all $ps < .001$). By contrast, Numeric Aggregation did not lead to significantly more additions or removals than Item Labels (additions: $t(1,191) = .57$,

$p = .567$, Cohen’s $d = .03$; removals: $t(1,148) = 1.14$, $p = .255$, Cohen’s $d = .07$). These results are robust to the inclusion of demographic factors (gender, age, BMI, income, numeracy, and dieting status; see Tables S3–S4 in the Web Appendix).

Mediation analyses. Reduced meal calories may be driven by meal revisions (operationalized as the number of item removals) that lead to the selection of items with fewer calories, or by guiding consumers to include fewer items in their final meal. We used mediation analysis to examine these relationships to better understand the choices behind the calorie change, comparing the two aggregation conditions while excluding the Item Labels condition. For all mediation analyses, we used the Numeric Aggregation condition as the baseline, and we standardized the variables of total calories, number of items ordered, and number of item removals. A bootstrapped mediation with 5,000 replications revealed that neither the number of items ordered ($\beta = -.018$, SE = .020, $z = .91$, $p = .36$) nor the number of item removals ($\beta = .009$, SE = .008, $z = 1.18$, $p = .24$) explained a significant proportion of the effect of Traffic Light Aggregation on meal calories for the full sample.

As a post hoc exploratory analysis, we investigated whether our predictions regarding the mediating role of revisions might be limited to those participants who needed guidance. Specifically, we reasoned that there should be little reason for additional revisions among participants ($n = 415$) who had already chosen a “green light” meal, whereas item removals (and possible substitutions) would be more likely among participants ($n = 785$) who had surpassed 750 calories in their meal and observed a yellow or red traffic light. To test this prediction, we estimated the same mediation model with a new dummy variable (“Surpassed 750”) included as a moderator of the relationship between Traffic Light Aggregation and number of item removals.

We again conducted a bootstrapped mediation with 5,000 replications, and in this moderated mediation model (see Figure 2, Panel D), among those who surpassed 750 calories at any point, the number of item removals explained a significant proportion of the effect of Traffic Light Aggregation on meal calories ($\beta = -.071$, SE = .016, 95% CI = [−.102, −.040], $z = 4.46$, $p < .001$), but not among those who never surpassed 750 calories ($\beta = -.0001$, SE = .004, $z = .01$, $p = .99$). In this model, the number of items ordered still did not explain a significant proportion of the effect of Traffic Light Aggregation on meal calories ($\beta = -.011$, SE = .012, $z = .90$, $p = .37$). In other words, the calorie-reducing effect of Traffic Light Aggregation is observed for those who receive a yellow or red light at some point during their meal selection. Among these participants, the effect was mediated by the tendency to make more revisions. In contrast, participants who consider only low-calorie meals show no difference in meal calorie content between the two aggregation conditions.

Discussion

Relative to the baseline, in which numeric calorie labels were presented only for individual items, both forms of aggregation led to

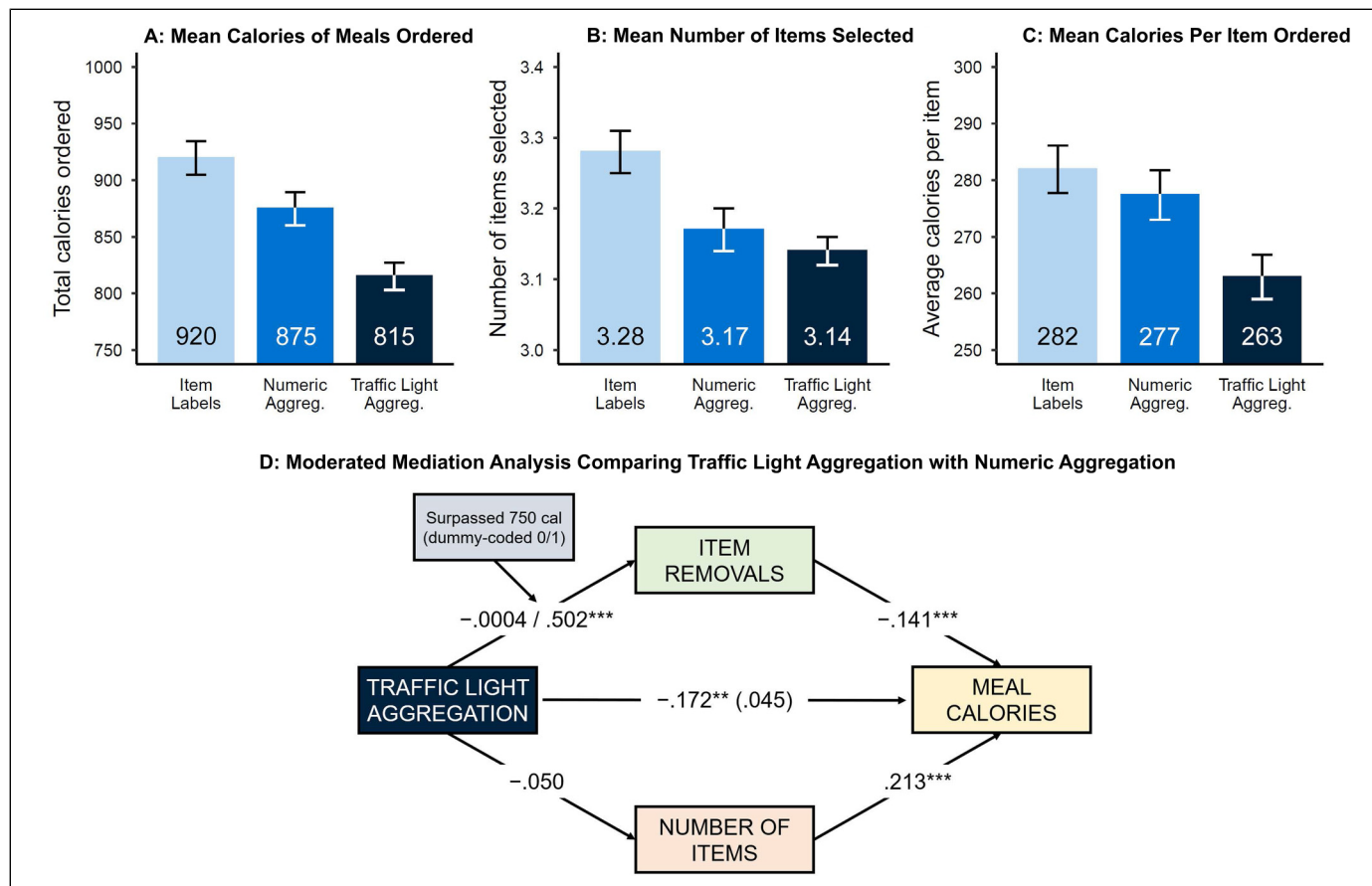


Figure 2. Experiment 2: Effect of traffic light aggregation (vs. item labels; vs. numeric aggregation) on meals ordered.

* $p < .05$. ** $p < .01$. *** $p < .001$.

Notes: Error bars in Panels A–C indicate ± 1 SE. Coefficients in Panel D are standardized betas. Aggreg. = Aggregation.

significant calorie reductions, with Traffic Light Aggregation further outperforming Numeric Aggregation. This added benefit appears to be driven by meal revisions among those seeing yellow or red lights, suggesting that the dynamically updating traffic lights led these consumers to consider more alternatives and ultimately decide on lower-calorie meals.

Overall, those presented with Traffic Light Aggregation ordered both fewer and lower-calorie items, leading to a 105-calorie reduction (-11.4%) in meal calories relative to item labels alone. However, the design of Experiment 2 does not allow us to distinguish between the prescriptive information conveyed by the traffic light thresholds (which could be mimicked by combining other labeling formats with direct prescriptive guidance) versus some other efficiency of traffic lights, such as their intuitive format. We address this concern in Experiments 3–5.

Experiment 3: Prescriptive Versus Intuitive Guidance

Because numeric information on its own lacks prescriptive information about the appropriateness of choices, research and policies implementing such labels have tried including numeric calorie reference guides alongside labels, such as

recommending 2,000 calories per day. Such guidelines do not seem to help consumers use item labels (Downs et al. 2013; Wisdom, Downs, and Loewenstein 2010), but they have not been tested with aggregated feedback. Experiment 3 extends the design of Experiment 2 by adding written numeric guidance for low-, moderate-, and high-calorie meals, corresponding to the same thresholds used for green, yellow, and red lights. However, our theorizing predicts that because this numeric guidance is less intuitive—because it requires interpretation of the numeric aggregation according to the categories identified—traffic light aggregation will continue to reduce calories ordered relative to numeric aggregation. In addition, Experiment 3 includes a control condition with no calorie labeling to allow for a quantitative benchmark of the effect sizes of both individual and aggregate labels.

Methods

Participants. We recruited 2,524 participants from MTurk to complete the online study. After excluding 87 participants (3%) who selected meals that cost over \$18 or contained over 2,000 calories (preregistered exclusion criteria), the final sample consisted of 2,437 participants (54% female, mean age = 35.4 years).

Procedure. Participants selected a meal from the same menu used in Experiment 2. Participants were again required to stay on the menu selection page for at least one minute and were told to imagine that they had a maximum budget of \$18. Participants were randomly assigned to one of four conditions: no labels (“Control”), numeric labels only on items (“Item Labels”), numeric item labels plus numeric aggregation (“Numeric Aggregation”), or numeric item labels plus traffic light aggregation (“Traffic Light Aggregation”). In all three labeled conditions, text at the bottom of the menu provided the following guidance: “Calorie content of meal: below 750 = relatively low; between 750 and 1,125 = moderate; over 1,125 = relatively high” (for the full list of items and menus, see Figures S6–S9 in the Web Appendix).

After selecting a meal, all participants indicated whether they noticed calorie guidelines at the bottom of the menu screen, provided their gender and age, and had the option to submit any final comments about the study. Over 90% of participants in each condition correctly answered whether calorie guidelines were present or absent—91% in the Control condition identified that calorie guidelines were absent, whereas 94% of Item Labels, 93% of Numeric Aggregation, and 95% of Light Aggregation participants correctly identified that the guidelines were present. We also conducted robustness checks on our main analyses by excluding participants who answered incorrectly. All results are robust to the exclusion of these participants as well as to the inclusion of demographic controls (see Tables S6–S11 in the Web Appendix).

Results

Calories ordered. Traffic Light Aggregation ($M = 858$ calories, 95% CI = [832, 884]) reduced calories compared with all other conditions, including Control ($M = 986$, 95% CI = [955, 1,018]; $t(1,177) = 6.21$, $p < .001$, Cohen’s $d = .36$), Item Labels alone ($M = 901$, 95% CI = [873, 929]; $t(1,208) = 2.22$, $p = .027$, Cohen’s $d = .13$), and Numeric Aggregation ($M = 903$, 95% CI = [874, 931]; $t(1,205) = 2.29$, $p = .022$, Cohen’s $d = .13$). The two numeric labeling conditions did not differ from each other ($t(1,214) = .09$, $p = .93$), but both Item Labels and Numeric Aggregation outperformed Control (Item Labels: $t(1,203) = 3.99$, $p < .001$, Cohen’s $d = .23$; Numeric Aggregation: $t(1,206) = 3.88$, $p < .001$, Cohen’s $d = .22$; see Figure 3, Panel A).

Number of items. Traffic Light Aggregation ($M = 3.23$ items, 95% CI = [3.18, 3.29]) did not reduce number of items relative to Item Labels ($M = 3.28$, 95% CI = [3.22, 3.34]; $p = .290$) and only marginally significantly reduced the number of items relative to Numeric Aggregation ($M = 3.32$, 95% CI = [3.25, 3.38]; $t(1,211) = 1.85$, $p = .064$, Cohen’s $d = .11$). All labeling conditions had fewer items than Control ($M = 3.43$, 95% CI = [3.36, 3.50]), including Item Labels ($t(1,192) = 2.97$, $p = .003$, Cohen’s $d = .17$), Numeric Aggregation ($t(1,187) = 2.29$, $p = .022$, Cohen’s $d = .13$), and Light Aggregation

($t(1,163) = 4.04$, $p < .001$, Cohen’s $d = .23$). There was no difference between Item Labels and Numeric Aggregation ($p = .45$; see Figure 3, Panel B).

Average calories per item. Traffic Light Aggregation ($M = 268$ calories, 95% CI = [261, 275]) did not significantly reduce calories per item relative to Numeric Aggregation ($M = 276$, 95% CI = [268, 284]; $p = .16$) and only marginally significantly reduced calories per item relative to Item Labels ($M = 278$, 95% CI = [270, 286]; $t(1,205) = 1.84$, $p = .066$, Cohen’s $d = .11$). Again, all labeling conditions reduced the average calories per item ordered relative to the Control condition ($M = 292$ calories, 95% CI = [283, 300]), including Item Labels ($t(1,213) = 2.24$, $p = .026$, Cohen’s $d = .13$), Numeric Aggregation ($t(1,216) = 2.58$, $p = .010$, Cohen’s $d = .15$), and Light Aggregation ($t(1,193) = 4.11$, $p < .001$, Cohen’s $d = .24$). There was no difference between Item Labels and Numeric Aggregation ($p = .70$; see Figure 3, Panel C).

Meal Revisions. Traffic Light Aggregation again led to significantly more meal revisions than the other conditions (for full details, including analyses for item additions, see the Web Appendix). Specifically, Traffic Light Aggregation caused more item removals ($M = 1.19$, 95% CI = [1.01, 1.36]) than Control ($M = .63$, 95% CI = [.54, .72]), Item Labels ($M = .65$, 95% CI = [.54, .77]), or Numeric Aggregation ($M = .79$, 95% CI = [.68, .90]) (all $ps < .001$). Numeric Aggregation also significantly increased removals compared with Control ($t(1,163) = 2.20$, $p = .028$, Cohen’s $d = .13$). There were no other significant differences between conditions in the number of item removals.

Mediation. We again conducted a moderated mediation analysis to test item removals and number of items, standardized, as possible mediators to explain the benefits of Traffic Light Aggregation relative to Numeric Aggregation (Figure 3, Panel D). A bootstrapped mediation with 5,000 replications revealed that the number of item removals explained a significant proportion of the effect of Traffic Light Aggregation on meal calories among those who surpassed 750 calories ($\beta = -.035$, SE = .011, 95% CI = [−.057, −.014], $z = 3.57$, $p = .001$), but not among those who did not surpass 750 calories ($\beta = .001$, SE = .004, $z = .39$, $p = .70$). Ordering more items also explained a marginally significant proportion of the effect of Traffic Light Aggregation ($\beta = -.034$, SE = .018, $z = 1.82$, $p = .068$).

Discussion

As in the previous studies, Traffic Light Aggregation reduced calories, mediated by more item removals among those who saw yellow or red lights while ordering. This calorie-reducing effect was significant even relative to a numeric condition that provided all the same information, including numeric guidance that mirrored the categories used by traffic lights, suggesting that the intuitive nature of the Traffic Light format has an additional impact in encouraging people to revise their meal. Interestingly, and in contrast to Experiment 2, Numeric

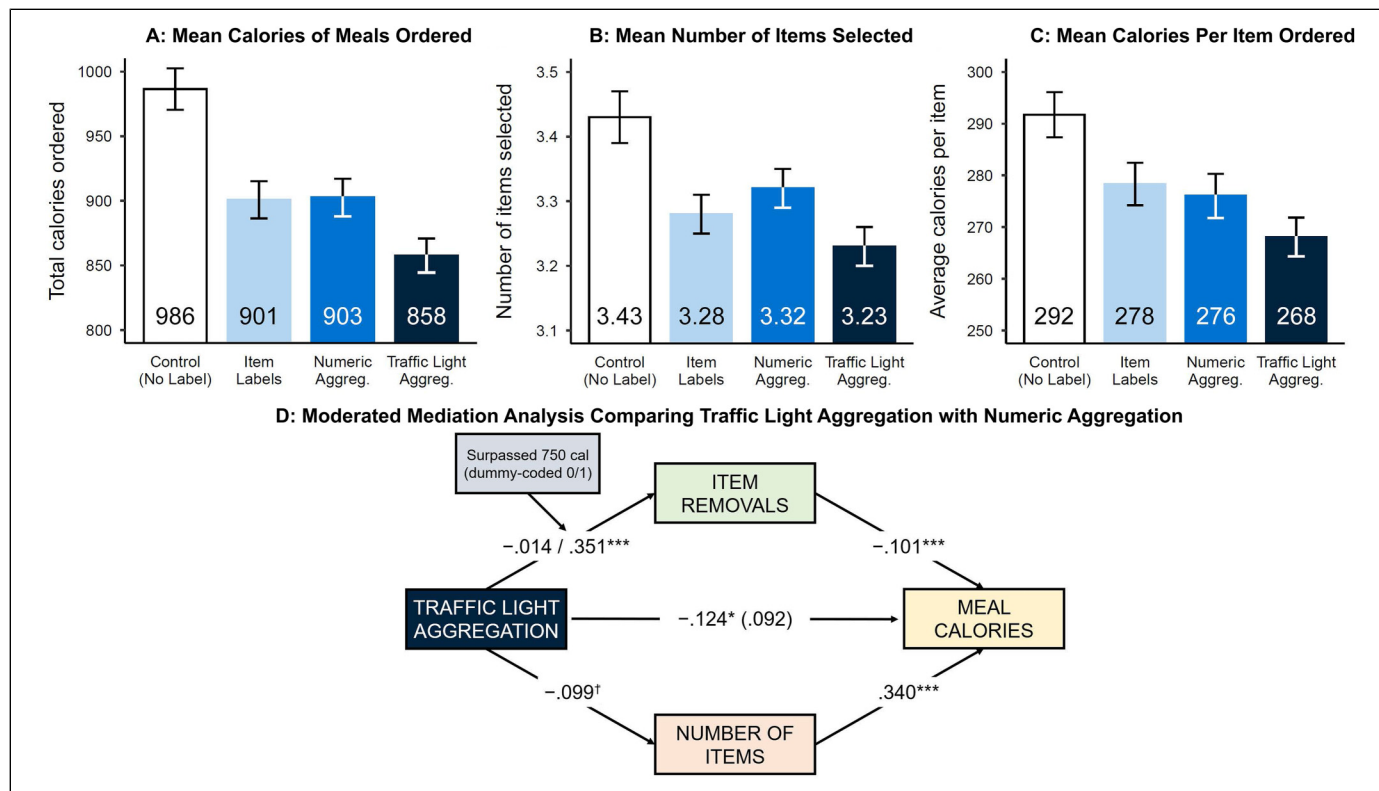


Figure 3. Experiment 3: Effect of traffic light aggregation (vs. control; vs. item labels; vs. numeric aggregation) on meals ordered.

[†] $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Notes: Error bars in Panels A–C indicate ± 1 SE. Coefficients in Panel D are standardized betas. Aggreg. = Aggregation.

Aggregation provided no additional benefit over item-only labels, despite the presence of numeric recommendations that corresponded to traffic light thresholds.

In combination, Experiments 1–3 reveal robust benefits of dynamically updated meal-level traffic light labels. They also show that the effectiveness of these labels comes from consumers' ability to add and remove items prior to finalizing an order when the meal-level traffic light changes color. An open question, however, is exactly what property of traffic lights yields these benefits. Traffic lights have two appealing properties that can guide meal revisions: (1) they intuitively map on to a specific action (e.g., a red light tells people to "stop" and change their behavior, whereas a green light tells people to "go ahead"), and (2) they provide discrete feedback that promotes action when a threshold is crossed. Which property is (most) critical for their observed benefits? The next experiments were designed to address this question.

Experiment 4: Discrete Categories or Intuitive Format?

To this point, we have argued that traffic light aggregation works by providing more actionable information than numeric information, guiding people to revise their meals in response to feedback that their meals have exceeded the calorie threshold where a green light turns into a yellow light. However, such a design feature of traffic light aggregation

potentially conflates two avenues by which this information provision affects revision decisions. Is the effect unique to the format of traffic light labels and their intuitive design that guides people to proceed with green light meals but revise their meals when they see a yellow or red light? Or could the same effect be found when any labeling approach provides discrete, categorical feedback with clear thresholds between low and moderate calorie categories? Employing a 2 (format: numeric vs. traffic light) $\times$ 2 (aggregation type: continuous vs. discrete) + 1 (no aggregation) design, Experiment 4 is aimed to disentangle these two features: feedback format and use of continuous versus categorical aggregation.

Methods

Participants. We recruited 4,008 participants from Prolific and excluded 1,005 participants (25%) who (1) selected meals that cost over \$18, did not order any items (violating instructions), or ordered over 2,000 calories ($n = 60$), or (2) failed the attention check question after the meal selection ($n = 946$).⁴ These exclusion criteria were preregistered. In addition,

⁴ All of the main results are robust to the inclusion of inattentive participants. We report detailed robustness checks (ordinary least squares regression analyses) that include the full sample of participants in Table S15 of the Web Appendix. In line with this robustness check, we do not include an attention check in subsequent studies.

we excluded two participants who did not select any items. The final sample contains 3,002 observations (52% female, mean age = 33.7 years).

Procedure. All participants chose a hypothetical meal from the same menu that we used in Experiments 2 and 3 (see Figures S10–S14 in the Web Appendix). Numeric calorie labels and prices were marked on each item, and participants were randomly assigned to one of five conditions: Item Labels (identical to Item Labels in Experiment 2), Continuous Numeric Aggregation (identical to Numeric Aggregation in Experiment 2), Discrete Numeric Aggregation, Continuous Traffic Light Aggregation, or Discrete Traffic Light Aggregation (identical to Traffic Light Aggregation in Experiment 2).

Participants in the Discrete Numeric Aggregation condition saw one of the following three messages on the top of the meal selection screen as they were making their decision: “Total Calories: below 750 cal,” “Total Calories: between 750 and 1,125 cal,” or “Total Calories: over 1,125 cal.” Thus, instead of displaying the running total of calorie content (as in the Continuous Numeric Aggregation condition), only categorical feedback was provided. These ranges were the same as the ranges that determined the green/yellow/red lights in the Discrete Traffic Light Aggregation condition (as in Experiments 2 and 3). To make the change between the numeric categories more salient, we made the feedback message move vertically (i.e., upward/downward when changing to higher/lower calorie categories), comparable to the vertical alignment of green, yellow, and red traffic lights.

Participants in the Continuous Traffic Light Aggregation condition saw a gradually changing traffic light as they were selecting their meal. Both the color of the traffic light and the vertical position of the light were programmed to be a continuous linear function of the selected total calories (i.e., there were no clear boundaries between green/yellow/red lights). The lower (higher) the total calorie content of the meal was, the greener (redder) the traffic light was, and the lower (higher) the light was positioned vertically. The color and the vertical position of the light in the Continuous Traffic Light Aggregation were identical to the “classic” colors and positions used in the Discrete Traffic Light condition at the midpoint of each range. That is, because the calorie thresholds for the meal-level traffic lights in the Discrete Traffic Light condition were 0–749 calories (green light), 750–1,124 calories (yellow light), and 1,125+ calories (red light), the traffic light in the Continuous Traffic Light aggregation was identical in color to these at 375 calories (green), 937.5 calories (yellow), and 1,312.5 calories (red). For values lower than 375 calories, we used the same green light as in the previous experiments, and for values greater than 1,312.5 calories, we used the same red light as before (see Figure 4, Panel A). As in the previous studies, each meal revision was recorded and time-stamped in our data. Following the meal selection, participants answered an attention check question, indicated whether they added or removed any item simply for the sake of exploration (“Did you add or remove any items just to explore how the menu works, out of curiosity?”; yes/no), and reported their gender and age.

Results

Calories ordered. Both versions of Traffic Light Aggregation (Discrete: $M = 871$ calories, 95% CI = [846, 896]; Continuous: $M = 860$, 95% CI = [835, 886]) significantly reduced total calories compared with Item Labels ($M = 956$, 95% CI = [925, 986]; Discrete: $t(1,159) = 4.24$, $p < .001$, Cohen’s $d = .24$; Continuous: $t(1,159) = 4.75$, $p < .001$, Cohen’s $d = .27$). We observed no significant difference between the two types of Traffic Light Aggregation ($t(1,200) = .57$, $p = .571$, Cohen’s $d = .03$). By contrast, Discrete Numeric Aggregation ($M = 928$, 95% CI = [899, 956]) and Continuous Numeric Aggregation ($M = 934$, 95% CI = [906, 963]) did not significantly reduce total calories compared with Item Labels (Discrete: $t(1,195) = 1.32$, $p = .188$, Cohen’s $d = .08$; Continuous: $t(1,196) = 1.00$, $p = .319$, Cohen’s $d = .06$). In addition, we found no significant difference in total calories ordered between the two types of Numeric Aggregation ($t(1,198) = .33$, $p = .743$, Cohen’s $d = .02$). Finally, both types of Traffic Light Aggregation reduced total calories significantly compared with both types of Numeric Aggregation (all $ps \leq .004$) (see Figure 4, Panel B).

Regression analyses. To investigate whether there was an interaction between aggregation type and feedback type, we conducted ordinary least squares regression analyses (Web Appendix, Tables S12–S16). Compared with Item Labels, Traffic Light Aggregation (either continuous or discrete) led to a significant reduction of total calories ($\beta = -95.3$, $t(2,997) = 4.77$, $p < .001$). Traffic Light Aggregation was also associated with a significant reduction in average calories per item ($\beta = -25.3$, $t(2,997) = 4.29$, $p < .001$), a significant increase in item additions ($\beta = .88$, $t(2,997) = 6.73$, $p < .001$), and a significant increase in item removals ($\beta = .96$, $t(2,997) = 7.97$, $p < .001$). By contrast, Numeric Aggregation did not differ from Item Labels in any way (all $ps > .10$).

There were no significant main effects of feedback type on total calories, items selected, or average calories per item, nor did we find a significant interaction between aggregation type and feedback type (all $ps > .10$). The only effect feedback type had was a slight increase in the number of revisions: when feedback was provided in a discrete format (either numeric or traffic light), people added marginally significantly more items ($\beta = .22$, $t(2,997) = 1.70$, $p = .089$) and removed significantly more items ($\beta = .24$, $t(2,997) = 1.97$, $p = .049$). All the main results are robust to demographic controls, the inclusion of participants who failed the attention check, or the exclusion of people who reported that they had added or removed items out of curiosity.

Discussion

Overall, those presented with Traffic Light Aggregation (either discrete or continuous) ordered lower-calorie items, leading to a 90-calorie reduction (–9.4%) in meal calories relative to labeling items alone. By contrast, Numeric Aggregation led to a

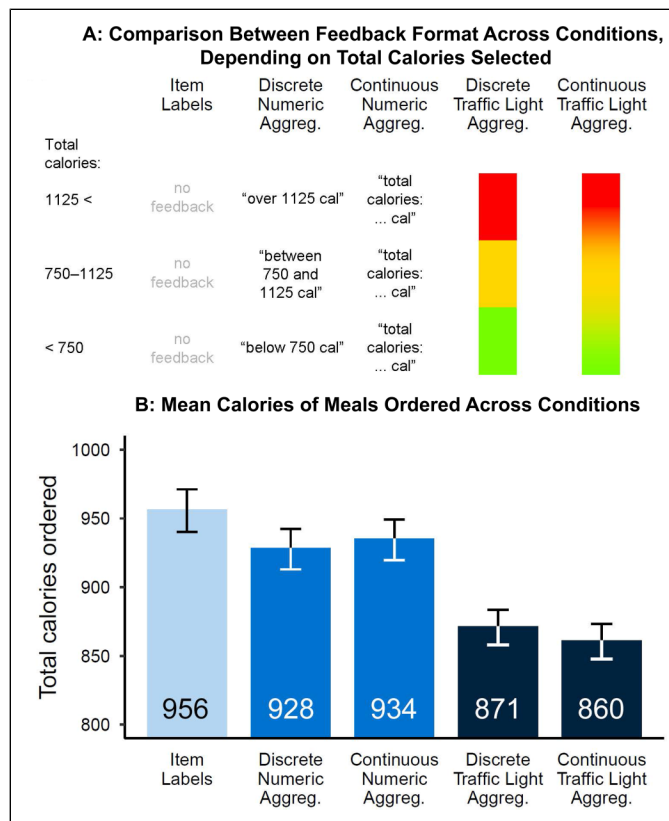


Figure 4. Experiment 4: Experimental manipulation and effects of labeling conditions on meal calorie content.

Notes: Error bars indicate ± 1 SE. For the actual meal selection screens in Panel A, see Figures S10–S14 in the Web Appendix. Aggreg. = Aggregation.

much smaller, nonsignificant 25-calorie reduction (-2.6%) relative to no aggregation. The lack of a main effect of feedback type and the lack of an interaction between feedback type and label type strongly suggest that it is not the categorical nature of the traffic light aggregation that prompts people to revise their order. Rather, it appears that the intuitive traffic light format triggers meal revisions, even when there are not discrete thresholds to discriminate between healthy and unhealthy meals. Thus, Experiment 4 demonstrates that gradual feedback presented as a continuously changing traffic light is as efficient as the classic three-color (discrete) traffic light feedback, and both of these are superior to numeric aggregation, regardless of how numeric feedback is provided.

Experiment 5: Intuitive Versus Unintuitive Feedback

In Experiment 5, we tested whether the effectiveness of the traffic light aggregation observed in previous studies was due to the intuitive nature of the traffic light graphic or to the increased visual salience offered by multicolored dynamic feedback. It is possible that dynamically updating numeric values do not grab as much attention as changing colors that represent the calorie content of the meal, and that any colorful or otherwise visually

salient display of total meal calories would work as well. By contrast, we predict that feedback can be very visually salient or use familiar symbols but will not achieve substantial calorie reductions unless its interpretation is intuitive. In Experiment 5, we compare our standard traffic light graphic with a less intuitive version of traffic lights with different colors, and we also test a new intuitive labeling system (emojis) against visually similar (but unintuitive) abstract symbols.

Methods

Participants. We recruited 3,036 participants from Prolific and excluded 26 participants (.9%) whose meals cost over \$18 or contained over 2,000 calories (exclusion criteria were preregistered). The final sample contains 3,010 observations (52% female, mean age = 31.8 years).

Procedure. Participants chose a hypothetical meal from the same menu that we used in Experiments 2–4 (see Figures S15–S19 in the Web Appendix). Meal selection was not time constrained: participants could spend as much (or as little) time selecting items from the menu as they wanted. Numeric calorie labels and prices were marked on each item, and participants were randomly assigned to one of five conditions: Item Labels (identical to Experiments 2 and 4), Intuitive Traffic Light Aggregation (identical to previous experiments), Unintuitive Traffic Light Aggregation, Intuitive Symbol Aggregation, or Unintuitive Symbol Aggregation.

Participants in the four aggregation conditions received real-time graphic feedback about the overall calorie content of their meals, based on the same thresholds used in Experiments 2–4), and we manipulated the type of graphic that represented the low/moderate/high ranges across conditions. In the Intuitive Traffic Light Aggregation condition, the feedback was presented in the form of the “classic” green/yellow/red traffic lights. In the Unintuitive Traffic Light Aggregation condition, participants saw a horizontally aligned traffic light graphic, with purple/white/blue indicating the low/moderate/high ranges of total calories.

In the Intuitive Symbol Aggregation condition, feedback was presented in the form of emojis (happy/neutral/sad representing low/moderate/high), which were chosen to be more intuitive in meaning than their control comparison, the Unintuitive Symbol Aggregation condition, consisting of equivalently colored yellow circles but with unintuitive symbols (#, @, and * representing low, moderate, and high, respectively) instead of faces. Even though these keyboard symbols are likely familiar to consumers, we argue that they are unintuitive in this context because their relationship to meal calories is unclear in the absence of an explanatory legend. By contrast, emoji feedback is intuitive because it clearly signals either happiness, neutrality, or sadness about the healthiness of the meal selected. Importantly, we kept the visual salience and size of the graphic feedback as similar as possible across conditions (see Figure 5, Panel A). All four aggregation conditions featured a prominent explanation

displayed directly under the graphic feedback to clarify which range each symbol or light represented. Thus, information was visually similar and semantically equivalent across conditions, but the unintuitive conditions were expected to require participants to consult the legend that explained what the lights meant, whereas the intuitive conditions were expected to require less translational work. As in the previous studies, each meal revision was recorded and time-stamped in our data. Following the meal selection, participants provided their gender and age to complete the study.

Results

Calories ordered. We report the effects of the different labeling conditions on meal calories in Figure 5, Panel B. Intuitive Traffic Light Aggregation ($M = 874$ calories, 95% CI = [850, 899]) significantly reduced total calories compared with Item Labels ($M = 970$, 95% CI = [941, 999]; $t(1,167) = 4.93$, $p < .001$, Cohen's $d = .28$). By contrast, Unintuitive Traffic Light Aggregation ($M = 940$, 95% CI = [914, 966]) did not significantly reduce total calories compared with Item Labels ($t(1,186) = 1.49$, $p = .137$, Cohen's $d = .09$). Indeed, Intuitive Traffic Light Aggregation led to significantly fewer calories than Unintuitive Traffic Light Aggregation ($t(1,198) = 3.63$, $p < .001$, Cohen's $d = .21$). This supports our prediction that the intuitiveness of traffic light aggregation drives the greatest reduction in meal calories, rather than mere novelty or salience.

As a further test of the benefits of intuitive labeling, we find that Intuitive Symbol (emoji) Aggregation ($M = 909$ calories, 95% CI = [881, 936]) significantly reduced total calories compared with Item Labels ($t(1,196) = 3.02$, $p = .003$, Cohen's $d = .17$), whereas the Unintuitive Symbol Aggregation ($M = 926$, 95% CI = [899, 954]) led to a more modest reduction of total calories compared with Item Labels ($t(1,198) = 2.13$, $p = .033$, Cohen's $d = .12$). That is, although graphic aggregation works to reduce meal calories, and even an arbitrary symbol like “#” can be combined with dynamic feedback to motivate calorie reductions, the largest reductions are consistently achieved when graphic aggregation uses intuitive imagery. Notably, Intuitive Traffic Light Aggregation ($M = 874$ calories) achieved the lowest-calorie meals of all conditions; it led to a marginally significant reduction of total calories compared with Intuitive Symbol Aggregation ($t(1,189) = 1.84$, $p = .066$, Cohen's $d = .11$) and statistically significant reductions of total calories compared with the Unintuitive Aggregation conditions (both $ps \leq .006$).

Discussion

Experiment 5 demonstrated that intuitive Traffic Light Aggregation (green/yellow/red lights) helps consumers reduce total calories primarily by providing intuitive feedback. This experiment refutes several alternative accounts that could explain the findings in previous studies. First, in Experiment 5, all aggregation conditions featured visually salient graphical feedback, yet we still observed superiority

of traffic light aggregation, which suggests that the results were not driven by mere enhanced visual salience. Second, we demonstrated that the effect of intuitive, dynamic feedback is not strictly limited to traffic lights, as we observed the intuitive emojis (happy, neutral, sad) to be almost as efficient as traffic lights. This, combined with the finding that the *unintuitive* traffic lights did not help consumers to reduce total calories, strongly suggests that it is the intuitiveness of the traffic light aggregation that primarily explains its effect, not necessarily its format (i.e., symbol vs. lights).

Internal Meta-Analysis of Effect Sizes Across All Experiments

To identify the average effect size of traffic light aggregation and to compare effect sizes across different labeling formats, we conducted a fixed-effects meta-analysis using the R packages *metafor* (Viechtbauer 2010) and *metaviz* (Kossmeier, Tran, and Voracek 2019). Using this approach, we determined the overall effect size (and associated standard error) of different calorie label formats, incorporating data from Experiments 1–5 as well as from Supplemental Experiments 1–4.⁵ To enable cross-study comparisons, we first standardized the total calories measure within each experiment and then calculated the standardized effect sizes (Cohen's d) for the differences between the numeric item labels only condition and the other conditions within each experiment. We chose the numeric item labels condition as a reference for our comparisons for two reasons: this condition was present in all studies, and this calorie labeling type is the most frequently used in real-world settings; therefore, it serves as the most informative and externally valid baseline. We estimated the average effect sizes (relative to numeric item labels) for five types of calorie labeling: (1) no label, (2) traffic light-item labels only (present only in supplemental experiments), (3) traffic light item labels combined with traffic light meal label (present only in supplemental experiments), (4) numeric item labels combined with numeric meal label (both discrete and continuous), and (5) numeric item labels combined with traffic light meal label (both discrete and continuous). We then calculated the overall standardized effects as the weighted average across experiments, using the study weights assigned by the meta-analysis. Figure 6 summarizes the results of this meta-analysis.

The overall effect of numeric item labels combined with a traffic light meal label was associated with the largest reduction in calories among all labeling types ($M = -.235$, 95% CI = $[-.275, -.194]$; $Z = 11.38$, $p < .001$). The overall effect of a numeric meal label combined with numeric item labels was also significant ($M = -.071$, 95% CI = $[-.122, -.020]$; $Z = 2.73$, $p = .006$) but was only about one-third as strong as the

⁵ We included only those conditions that were present in at least two experiments in the meta-analysis. Therefore, we excluded three Experiment 5 conditions—Unintuitive Traffic Light Aggregation, Unintuitive Symbol Aggregation, and Intuitive Symbol Aggregation—as there was no overall effect size to calculate using this method.

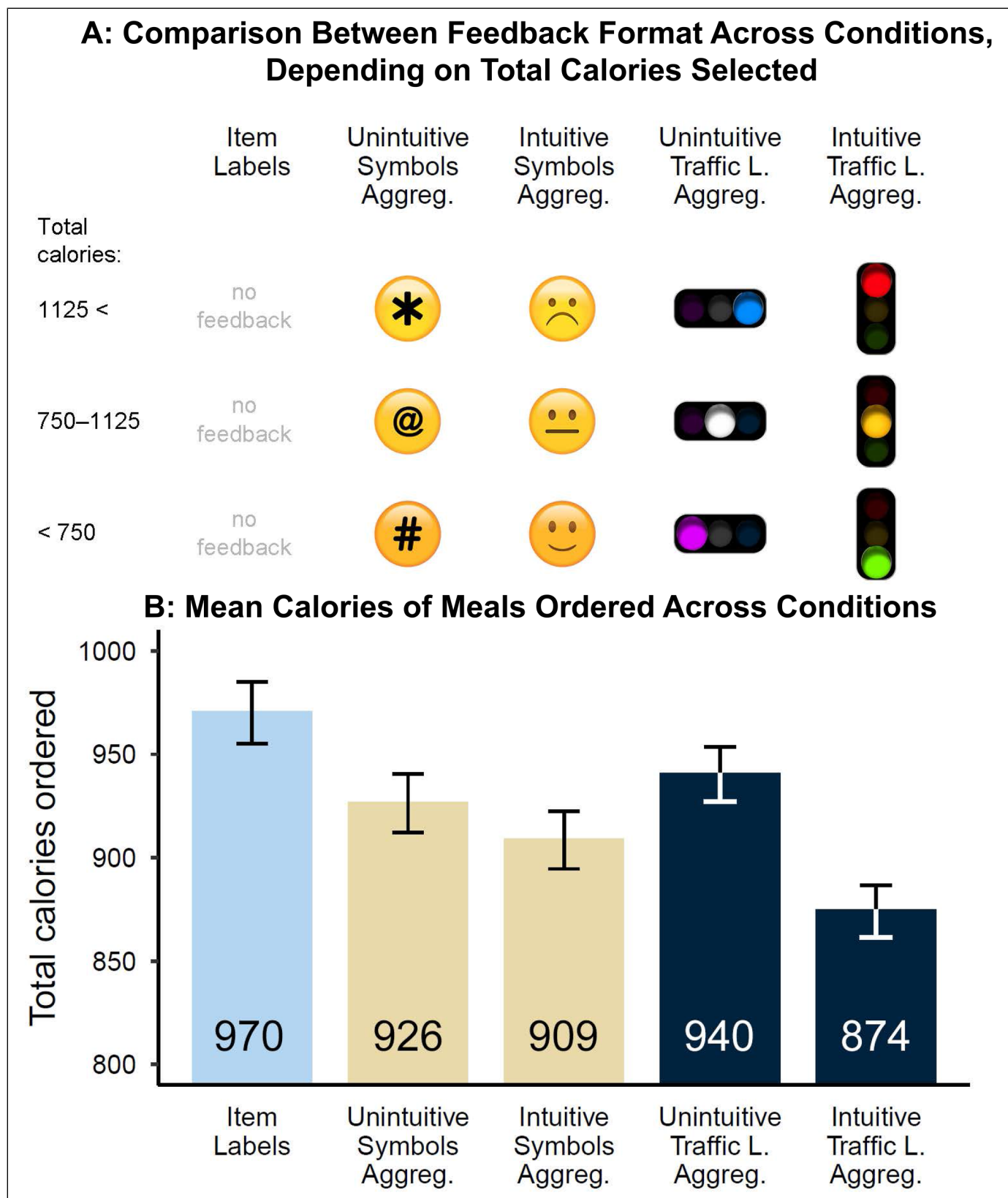


Figure 5. Experiment 5: Experimental manipulation and effects of labeling conditions on meal calorie content. Notes: Error bars indicate ± 1 SE. Traffic L. = Traffic Light; Aggreg. = Aggregation.

effect of traffic light aggregation. When only traffic light item labels were present, there was no significant reduction in calories compared with numeric item labels ($M = -.031$, 95% CI = $[-.135, .074]$; $Z = .57$, $p = .568$). When traffic light item labels were combined with a traffic light meal label, we found a significant reduction of calories ($M = -.121$, 95% CI = $[-.224, -.018]$; $Z = 2.29$, $p = .022$), though this effect was only about half as strong as the effect of traffic light aggregation combined with numeric item labels. Finally, the effect of the absence of calorie labeling was significantly positive; that is, participants ordered significantly more calories when they had no calorie information compared with when they were exposed to item labels ($M = .121$, 95% CI = $[.045, .192]$; $Z = 3.29$, $p = .001$).

General Discussion

Across five experiments, we demonstrate the benefits of an intuitive meal-level calorie label. Specifically, dynamically aggregated meal-level calorie labels presented on menus in a traffic light format reduce total calories ordered. An internal meta-analysis of these experiments showed that traffic light aggregation is the most effective labeling approach, leading to a reduction in meal calories considerably larger than numeric aggregation or item labels alone. Traffic light aggregation outperforms other labeling formats even when recommendations for meal calorie content are provided, and regardless of whether aggregated feedback is continuous or presented in discrete categories. This suggests that the intuitive nature of the traffic light format is responsible for its superior effectiveness. Although real-time aggregated calorie information reduces the number of items ordered across all labeling formats, including numeric aggregation, only traffic light aggregation leads to additional calorie reductions by encouraging people to replace high-calorie items with healthier substitutes.

We also show consistent and robust evidence for the process by which traffic light aggregation leads to calorie reductions relative to numeric aggregation: among those who consider higher-calorie meals, warnings provided by traffic light aggregation (i.e., the appearance of a yellow or red light) prompt more revisions, which lead to selections of lower-calorie meals.

One important question about the preceding studies is whether providing meal-level traffic light labels is the optimal approach, or whether alternative labeling (e.g., featuring static traffic light labeling on individual items) would work equally well. In a pair of supplemental experiments conducted across two different populations (Web Appendix, Supplemental Experiments 2 and 3), we tested a range of labeling approaches, including the use of traffic light calorie labels for individual items instead of numeric item labels. In these experiments, we again found that providing dynamically updated meal-level calorie information reduced the calorie content of orders relative to static information.

We also conducted an experiment (Web Appendix, Supplemental Experiment 4) in which we found feedback interventions to work similarly well even in the presence of

increased cognitive load. While paying attention to an audio conversation, consumers seeing traffic light aggregation continued to order fewer calories compared with those who only had numeric item labels. This seems to indicate that rather than simply motivating consumers to think harder about their selections (as prescriptive guidance might achieve), traffic light aggregation facilitates lower-calorie decisions in an intuitive manner that does not require much working memory.

Implications for Marketers and Managers

The COVID-19 pandemic and corresponding restrictions on in-person dining options have forced many food marketers to consider ways to engage their consumers via virtual platforms, reducing the direct contact between diners and restaurant staff. The delivery apps DoorDash, UberEats, Grubhub, and Postmates saw their collective revenue more than double over a six-month period in 2020, but consumer retention by individual apps has remained quite low, suggesting a need for additional ways to appeal to customers and build brand loyalty given the expenses associated with delivery and marketing (Dhillon and Wu 2021). Online food ordering platforms can clearly leverage these findings to promote healthier meal selections, attract health-conscious diners, and even increase customer satisfaction (see Supplemental Experiment 1 in the Web Appendix). Furthermore, the technological advances that allowed this study to deliver dynamic aggregation of information are not restricted to delivery services. Self-ordering kiosks are now available in many restaurants and food courts, mobile ordering for curbside pickup has experienced a substantial uptick in popularity, and cafeterias (e.g., at hospitals or universities) have frequently shifted meal ordering to virtual environments, where they can then use aggregated labeling to help their customers make healthier decisions. The benefits of providing consumers with aggregated nutrition information in real time, and allowing them to adjust their decisions prior to final purchasing decisions, can be delivered using a diverse array of digital tools such as websites, mobile apps, and tabletop ordering tablets (Gao and Su 2018).

We believe restaurants can leverage this tool to simultaneously promote themselves as a healthy brand and enhance customer satisfaction and sales. Though it is possible that some restaurant patrons would resist prescriptive efforts to influence their ordering decisions, we propose that a dynamic feedback system feels more flexible and responsive (and less paternalistic) than other efforts to influence health decisions and should meet with less consumer reactance or political opposition as a result. Rather than identifying particular items as healthy or unhealthy, or trying to set healthy defaults (e.g., a turkey burger rather than beef burger) that may backfire by driving customers away (Colby, Li, and Chapman 2020), menus that provide guidance at the meal level via aggregated feedback encourage customers to select balanced meals and facilitate complementary choices (Dhar and Simonson 1999). A meal can include a dessert or soda and still receive a green or yellow light, so long as consumers use the information provided to choose a low-calorie entree. Restaurants could suggest meal combinations of “vice-virtue bundles” (Liu, Haws,

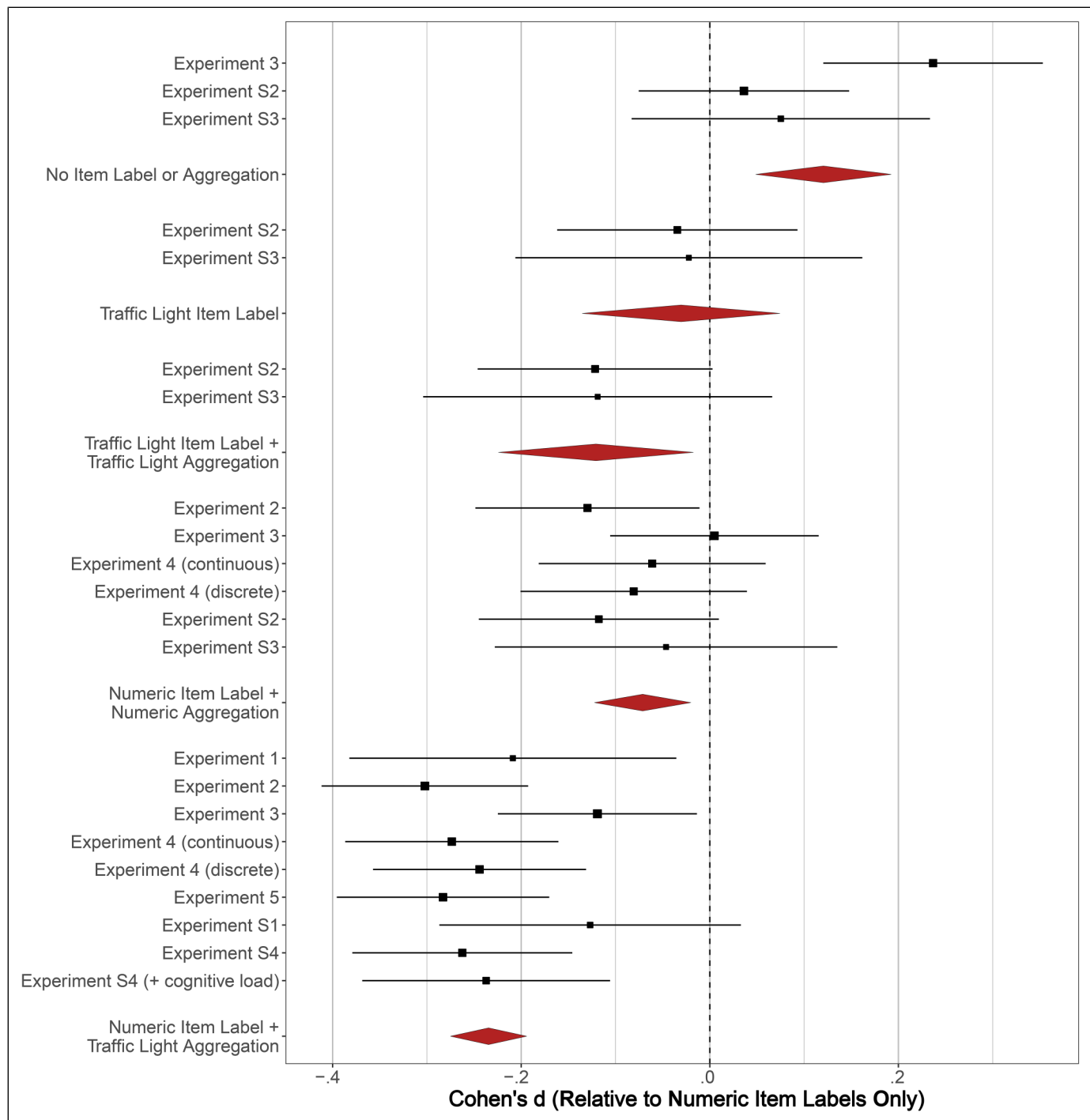


Figure 6. Results of the fixed-effects meta-analysis, dependent measure: total calories.

Notes: Standardized effect sizes are relative to the numeric item label condition within each experiment. Experiment 3 also featured a numeric guideline in all labeled conditions.

et al. 2015) or allow consumers to build their own combinations with varying sizes of relevant options (Haws and Liu 2016), incorporating both virtuous options (e.g., an entrée-sized salad) and unhealthier vice options (e.g., a small side order of French fries) in a single bundle that receives a green-light label as a meal.

Restaurants and food delivery services may need to be creative in scaling aggregation to decisions other than individual

meals; when consumers go online to order meals for multiple people, websites can facilitate the partitioning of orders by asking for a name for each meal—a practice that may in turn help the restaurant to organize and label orders for pickup. Feedback becomes more difficult when items are meant to be shared (like pizzas), though this limitation applies to item-level labeling as well. Websites could offer consumers the ability to

create personalized profiles to tailor traffic light labels based on consumers' height, weight, or even individual preferences regarding the size of a given meal (e.g., adjusted calorie thresholds for an afternoon snack or a pizza to be shared). Such personalized feedback need not be restricted to calories but could be customized according to nutritional information that is closely aligned with other aspects of one's medical history or goals, such as basing labels on sodium content for those with hypertension. Labeling could even go beyond nutritional information, such as providing information about the environmental impact of food items (Camilleri et al. 2019).

Future research could test the impact of different traffic light thresholds to fine-tune a traffic light guidance approach; our Experiment 1 used different calorie thresholds than Experiments 2–5 but found similar effects; Experiment 4 showed that even continuous traffic light labels that remove defined “thresholds” similarly promote calorie reductions. It remains open to investigation whether there is an ideal calorie content to equate to each color light, whether these calorie-color associations should vary between consumers, what thresholds should be used for non-calorie-based guidance, and how widely color-coding could vary and still obtain the desired outcome. Going forward, researchers could also extend these principles of intuitive, dynamic feedback to other decision-making contexts, such as shopping while on a financial budget or planning an organization's expenditures for the upcoming quarter.

Applications Beyond Food Labeling

Intuitive labels, such as the traffic light format used here, have applications beyond calorie information and food ordering. There is a wide range of situations in which the long-term interests of individuals require them to stop when some limit is reached or reconsider their actions when some threshold is passed (e.g., shopping, gambling, staying up too late, spending time on social media). Individuals often choose to avoid information that would interfere with their intuitive preferences (Woolley and Risen 2018) or their desire to yield to the immediate gratification offered by tempting options (Bazerman, Tenbrunsel, and Wade-Benzoni 1998; Bitterly et al. 2015). However, we show that traffic light labels provided as real-time feedback promote revisions of initial choices. Particularly when people either lack insight into the long-term consequences of their own actions or have problems with willpower, traffic lights provide the kind of useful, clear signal that can potentially change behavior for the better. Although the literature on interventions to promote self-control is already quite vast (for a recent review, see Duckworth, Milkman, and Laibson [2018]), dynamic traffic light labeling may be a novel and particularly flexible approach to aid individuals in many different situations.

With the advent of the internet and smartphones, there are expanded opportunities for providing people with *aggregated* information as well. For example, smartphone users can learn their screen time over the past week, both cumulative and per application, perhaps motivating a shift away from an addictive game or social media platform. Similarly, pedometers actively

running in the background of one's phone can provide aggregated step count information over the day, week, month, or year to encourage exercise goals and identify patterns of good and bad performance. Real-time aggregation of information regarding others' behavior can also motivate action—for example, learning that others have recently donated to a charitable cause can make it seem more likely that the charity's goal will be reached, making one's contributions feel important and likely to have an immediate impact (Camilleri and Larrick 2019).

Even more broadly, the current research highlights the potential benefit from providing people with more intuitive feedback in a wide range of situations. As is the case for traffic light labeling, this can involve providing feedback using symbols and systems that people can immediately comprehend in contexts that people would otherwise have difficulty understanding and navigating. Such labeling systems can even be customized to match the spirit of the underlying information. For example, if people at risk of hypertension do not have an intuitive understanding of blood pressure, they could potentially be given feedback in the form of an inflating balloon rather than (or in conjunction with) blood pressure numbers, thus conveying the essential meaning that as blood pressure (balloon inflation) increases, danger increases as well. For people who need help maintaining a pool of savings for incidental or emergency expenses, instead of providing numerical feedback about how much they have saved, banks could provide a picture of a reservoir that is more or less empty or full. These types of feedback might also be conducive to gamification—for instance, the saver could be presented with dynamic information in which their goal is to maintain the depth of water in a reservoir, with game-like warnings and milestones presented to add a sense of urgency so people do not delay in adding funds.

The ability to examine consumers' dynamic choices—their addition and removal of items—exploited in our analyses has potential benefits that go beyond research. Retailers and marketing researchers are always searching for information to differentiate consumers into “types” with unique patterns of behavior who may respond differently to specific environments and interventions. Consumers' dynamic decisions in the context of food, or indeed any domain, have the potential to provide such ancillary path data beyond their ultimate choices (Hui, Fader, and Bradlow 2009). Many websites already collect, and presumably make use of, not only consumers' clicks on websites but even their mouse movements (Englehardt 2017). Our research suggests that consumer behavior in response to dynamic feedback, such as that provided by traffic light labels, may be especially diagnostic.

Concluding Thoughts

Traffic light calorie aggregation to provide a dynamic, cumulative signal regarding the healthiness of a meal helps consumers order lower-calorie meals by encouraging revisions of meals in real time. These results have clear policy implications to expand traffic light calorie labeling to promote lower calorie consumption, which should be considered in the context of industry and

consumer responses to individual item labeling. Given persistent consumer demand for tools to help make healthier decisions, traffic light aggregation may be a meaningful and exciting offering for restaurants, cafeterias, websites, and apps to consider. Decision makers stand to benefit from platforms that provide them with not merely more information but timely, actionable information that identifies when to go ahead with a decision as planned and when they need to stop and reconsider.

Acknowledgment

The authors give special thanks to Dan Wall for his helpful comments in the design of Experiment 4.

Authors' Note

Eric M. VanEpps and Andras Molnar are joint first authors.

Associate Editor

James Bettman

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research was supported by EMV's and GL's personal research funds.

ORCID iD

Eric M. VanEpps  <https://orcid.org/0000-0002-9454-4847>

References

- Auchincloss, Amy H., Giridhar G. Mallya, Beth L. Leonberg, Andrew Ricchezza, Karen Glanz, and Donald F. Schwarz (2013), "Customer Responses to Mandatory Menu Labeling at Full Service Restaurants," *American Journal of Preventive Medicine*, 45 (6), 710–19.
- Bazerman, Max H., Ann E. Tenbrunsel, and Kimberly Wade-Benzoni (1998), "Negotiating with Yourself and Losing: Making Decisions with Competing Internal Preferences," *Academy of Management Review*, 23 (2), 225–41.
- Bettman, James R., Mary Frances Luce, and John W. Payne (1998), "Constructive Consumer Choice Processes," *Journal of Consumer Research*, 25 (3), 187–217.
- Bitterly, T. Bradford, Robert Mislavsky, Hengchen Dai, and Katherine L. Milkman (2015), "Dueling with Desire: A Synthesis of Past Research on Want/Should Conflict," in *The Psychology of Desire*, W. Hofmann and L.F. Nordgren, eds. New York: Guilford Press, 244–65.
- Bleich, Sara N., Christina D. Economos, Marie L. Spiker, Kelsey A. Vercammen, Eric M. VanEpps, Jason P. Block, et al. (2017), "A Systematic Review of Calorie Labeling and Modified Calorie Labeling Interventions: Impact on Consumer and Restaurant Behavior," *Obesity*, 25 (12), 2018–44.
- Block, Jason P. (2018), "The Calorie-Labeling Saga—Federal Preemption and Delayed Implementation of Public Health Law," *New England Journal of Medicine*, 379 (2), 103–05.
- Block, Jason P., Suzanne K. Condon, Ken Kleinman, Jewel Mullen, Stephanie Linakis, Sheryl Rifas-Shiman, et al. (2013), "Consumers' Estimation of Calorie Content at Fast Food Restaurants: Cross Sectional Observational Study," *BMJ*, 346, f2907.
- Bollinger, Bryan, Phillip Leslie, and Alan Sorensen (2011), "Calorie Posting in Chain Restaurants," *American Economic Journal: Economic Policy*, 3 (1), 91–128.
- Burton, Scot, Laurel Aynne Cook, Elizabeth Howlett, and Christopher L. Newman (2015), "Broken Halos and Shattered Horns: Overcoming the Biasing Effects of Prior Expectations Through Objective Information Disclosure," *Journal of the Academy of Marketing Science*, 43 (2), 240–56.
- Cadario, Romain and Pierre Chandon (2020), "Which Healthy Eating Nudges Work Best? A Meta-Analysis of Field Experiments," *Marketing Science*, 39 (3), 465–86.
- Camilleri, Adrian R. and Richard P. Larrick (2019), "The Collective Aggregation Effect: Aggregating Potential Collective Action Increases Prosocial Behavior," *Journal of Experimental Psychology: General*, 148 (3), 550–69.
- Camilleri, Adrian R., Richard P. Larrick, Shajuti Hossain, and Dalia Patino-Echeverri (2019), "Consumers Underestimate the Emissions Associated with Food but Are Aided by Labels," *Nature Climate Change*, 9 (1), 53–58.
- Cantor, Jonathan, Alejandro Torres, Courtney Abrams, and Brian Elbel (2015), "Five Years Later: Awareness of New York City's Calorie Labels Declined, with No Changes in Calories Purchased," *Health Affairs*, 34 (11), 1893–1900.
- Choi, Jungsil, Yexin Jessica Li, and Adriana Samper (2019), "The Influence of Health Motivation and Calorie Ending on Preferences for Indulgent Foods," *Journal of Consumer Research*, 46 (3), 606–19.
- Cohen, Deborah, Rajiv Bhatia, Mary T. Story, Margo Wootan, Christina D. Economos, and Linda Van Horn, et al. (2013), *Performance Standards for Restaurants: A New Approach to Addressing the Obesity Epidemic*. Santa Monica, CA: Rand Corporation.
- Colby, Helen, Meng Li, and Gretchen Chapman (2020), "Dodging Dietary Defaults: Choosing Away from Healthy Nudges," *Organizational Behavior and Human Decision Processes*, 161, 50–60.
- Dhar, Ravi and Itamar Simonson (1999), "Making Complementary Choices in Consumption Episodes: Highlighting Versus Balancing," *Journal of Marketing Research*, 36 (1), 29–44.
- Dhillon, Sunny and Kevin Wu (2021), "Delivery 2.0: How On-Demand Meal Services Will Become Something Far Bigger," *Fast Company* (February 15), <https://www.fastcompany.com/90604082/future-of-on-demand-meal-delivery-ghost-kitchens-postmates-doordash-uber-eats>.
- Downs, Julie S., Jessica Wisdom, and George Loewenstein (2015), "Helping Consumers Use Nutrition Information: Effects of Format and Presentation," *American Journal of Health Economics*, 1 (3), 326–44.
- Downs, Julie S., Jessica Wisdom, Brian Wansink, and George Loewenstein (2013), "Supplementing Menu Labeling with Calorie Recommendations to Test for Facilitation Effects," *American Journal of Public Health*, 103 (9), 1604–09.

- Duckworth, Angela L., Katherine L. Milkman, and David Laibson (2018), "Beyond Willpower: Strategies for Reducing Failures of Self-Control," *Psychological Science in the Public Interest*, 19 (3), 102–29.
- Elbel, Brian (2011), "Consumer Estimation of Recommended and Actual Calories at Fast Food Restaurants," *Obesity*, 19(10), 1971–78.
- Elbel, Brian, Rogan Kersh, Victoria L. Brescoll, and L. Beth Dixon (2009), "Calorie Labeling and Food Choices: A First Look at the Effects on Low-Income People in New York City," *Health Affairs*, 28 (6), w1110–21.
- Elbel, Brian, Tod Mijanovich, L. Beth Dixon, Courtney Abrams, Beth Weitzman, Rogan Kersh, et al. (2013), "Calorie Labeling, Fast Food Purchasing and Restaurant Visits," *Obesity*, 21 (11), 2172–79.
- Ellison, Brenna, Jayson L. Lusk, and David Davis (2013), "Looking at the Label and Beyond: The Effects of Calorie Labels, Health Consciousness, and Demographics on Caloric Intake in Restaurants," *International Journal of Behavioral Nutrition and Physical Activity*, 10 (1), 21.
- Ellison, Brenna, Jayson L. Lusk, and David Davis (2014), "The Impact of Restaurant Calorie Labels on Food Choice: Results from a Field Experiment," *Economic Inquiry*, 52 (2), 666–81.
- Engelhardt, Steven (2017), "No Boundaries: Exfiltration of Personal Data by Session-Replay Scripts," *Freedom to Tinker* (November 15), <https://freedom-to-tinker.com/2017/11/15/noboundaries-exfiltration-of-personal-data-by-session-replay-scripts/>.
- Federal Register (2014), "Food Labeling; Nutrition Labeling of Standard Menu Items in Restaurants and Similar Retail Food Establishments," (accessed June 24, 2020), <https://www.federalregister.gov/documents/2014/12/01/2014-27833/food-labeling-nutritionlabeling-of-standard-menu-items-in-restaurants-and-similar-retail-food>.
- Finkelstein, Eric A., Kiersten L. Strombotne, Nadine L. Chan, and James Krieger (2011), "Mandatory Menu Labeling in One Fast-Food Chain in King County, Washington," *American Journal of Preventive Medicine*, 40 (2), 122–27.
- Fishbach, Ayelet and Ravi Dhar (2005), "Goals as Excuses or Guides: The Liberating Effect of Perceived Goal Progress on Choice," *Journal of Consumer Research*, 32 (3), 370–77.
- Franckle, Rebecca L., Jason P. Block, and Christina A. Roberto (2016), "Calorie Underestimation When Buying High-Calorie Beverages in Fast-Food Contexts," *American Journal of Public Health*, 106 (7), 1254–55.
- Gao, Fei and Xuanming Su (2018), "Omnichannel Service Operations with Online and Offline Self-Order Technologies," *Management Science*, 64 (8), 3595–3608.
- Goswami, Indranil and Oleg Urminsky (2019), "The 'Mere Reminder' Effect of Salient Calorie Labeling," working paper, <https://doi.org/10.31234/osf.io/a4y8f>.
- Gustafson, Christopher R. and Eliana Zeballos (2019), "Cognitive Aids and Food Choice: Real Time Calorie Counters Reduce Calories Ordered and Correct Biases in Calorie Estimates," *Appetite*, 141, 104320.
- Haws, Kelly L. and Peggy J. Liu (2016), "Combining Food Type(s) and Food Quantity Choice in a New Food Choice Paradigm Based on Vice-Virtue Bundles," *Appetite*, 103, 441–49.
- Howlett, Elizabeth A., Scot Burton, Kenneth Bates, and Kyle Huggins (2009), "Coming to a Restaurant Near You? Potential Consumer Responses to Nutrition Information Disclosure on Menus," *Journal of Consumer Research*, 36 (3), 494–503.
- Hui, Sam K., Peter S. Fader, and Eric T. Bradlow (2009), "Path Data in Marketing: An Integrative Framework and Prospectus for Model Building," *Marketing Science*, 28 (2), 320–35.
- Khan, Uzma and Ravi Dhar (2006), "Licensing Effect in Consumer Choice," *Journal of Marketing Research*, 43 (2), 259–66.
- Kossmeier, Michael, Ulrich S. Tran, and Martin Voracek (2019) "Visualizing Meta-Analytic Data with R Package Metaviz," R package version 0.3.0, <https://CRAN.R-project.org/package=metaviz>.
- Larrick, Richard P., Jack B. Soll, and Ralph L. Keeney (2015), "Designing Better Energy Metrics for Consumers," *Behavioral Science & Policy*, 1 (1), 63–75.
- Liu, Peggy J., James R. Bettman, Arianna R. Uhalde, and Peter A. Ubel (2015), "'How Many Calories Are in My Burrito?' Improving Consumers' Understanding of Energy (Calorie) Range Information," *Public Health Nutrition*, 18 (1), 15–24.
- Liu, Peggy J., Kelly L. Haws, Cait Lamberton, 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.
- Parker, Jeffrey R. and Donald R. Lehmann (2014), "How and When Grouping Low-Calorie Options Reduces the Benefits of Providing Dish-Specific Calorie Information," *Journal of Consumer Research*, 41 (1), 213–35.
- Payne, John W., James R. Bettman, and David A. Schkade (1999), "Measuring Constructed Preferences: Towards a Building Code," *Journal of Risk and Uncertainty*, 19 (1), 243–70.
- Schwartz, Janet, Jason Riis, Brian Elbel, and Dan Ariely (2012), "Inviting Consumers to Downsize Fast-Food Portions Significantly Reduces Calorie Consumption," *Health Affairs*, 31 (2), 399–407.
- Shah, Avni M., James R. Bettman, Peter A. Ubel, Punam Anand Keller, and Julie A. Edell (2014), "Surcharges Plus Unhealthy Labels Reduce Demand for Unhealthy Menu Items," *Journal of Marketing Research*, 51 (6), 773–89.
- Sheehan, Daniel and Koert Van Ittersum (2018), "In-Store Spending Dynamics: How Budgets Invert Relative-Spending Patterns," *Journal of Consumer Research*, 45 (1), 49–67.
- Simmons, Joseph P., Leif D. Nelson, and Uri Simonsohn (2012), "A 21-Word Solution," *SSRN*, <http://dx.doi.org/10.2139/ssrn.2160588>
- Sonnenberg, Lillian, Emily Gelsomin, Douglas E. Levy, Jason Riis, Susan Barraclough, and Anne N. Thorndike (2013), "A Traffic Light Food Labeling Intervention Increases Consumer Awareness of Health and Healthy Choices at the Point-of-Purchase," *Preventive Medicine*, 57 (4), 253–57.
- Taksler, Glen B. and Brian Elbel (2014), "Calorie Labeling and Consumer Estimation of Calories Purchased," *International Journal of Behavioral Nutrition and Physical Activity*, 11 (1), 1–8.

- Tangari, Andrea Heintz, My Bui, Kelly L. Haws, and Peggy J. Liu (2019), "That's Not So Bad, I'll Eat More! Backfire Effects of Calories-Per-Serving Information on Snack Consumption," *Journal of Marketing*, 83 (1), 133–50.
- Thorndike, Anne N., Jason Riis, Lillian M. Sonnenberg, and Douglas E. Levy (2014), "Traffic-Light Labels and Choice Architecture: Promoting Healthy Food Choices," *American Journal of Preventive Medicine*, 46 (2), 143–49.
- Thorndike, Anne N., Lillian M. Sonnenberg, Jason Riis, Susan Barraclough, and Douglas E. Levy (2012), "A 2-Phase Labeling and Choice Architecture Intervention to Improve Healthy Food and Beverage Choices," *American Journal of Public Health*, 102 (3), 527–33.
- Ungemach, Christoph, Adrian R. Camilleri, Eric J. Johnson, Richard P. Larrick, and Elke U. Weber (2018), "Translated Attributes as Choice Architecture: Aligning Objectives and Choices Through Decision Signposts," *Management Science*, 64 (5), 2445–59.
- VanEpps, Eric M., Julie S. Downs, and George Loewenstein (2016), "Calorie Label Formats: Using Numeric and Traffic Light Calorie Labels to Reduce Lunch Calories," *Journal of Public Policy & Marketing*, 35 (1), 26–36.
- Viechtbauer, Wolfgang (2010), "Conducting Meta-Analyses in R with the Metafor Package," *Journal of Statistical Software*, 36 (3), 1–48.
- Wilcox, Keith, Beth Vallen, Lauren Block, and Gavan J. Fitzsimons (2009), "Vicarious Goal Fulfillment: When the Mere Presence of a Healthy Option Leads to an Ironically Indulgent Decision," *Journal of Consumer Research*, 36 (3), 380–93.
- Wisdom, Jessica, Julie S. Downs, and George Loewenstein (2010), "Promoting Healthy Choices: Information Versus Convenience," *American Economic Journal: Applied Economics*, 2 (2), 164–78.
- 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), 230–45.