Pretty Healthy Food: How and When Aesthetics Enhance Perceived Healthiness

Linda Hagen

University of Southern California

(Forthcoming at *Journal of Marketing*;
post-print, i.e., **not** the version of record)

Author Note

Linda Hagen (linda.hagen@marshall.usc.edu) is assistant professor of marketing at the Marshall School of Business, University of Southern California, Los Angeles, CA 90089. The author thanks Ed O'Brien, Brent McFerran, Kristin Diehl, and Kirk Kristofferson for helpful comments on earlier versions of this manuscript, as well as Elisa Solinas for assistance collecting data for study 2. Supplementary materials are included in the Web Appendix accompanying this manuscript.

Pretty Healthy Food: How and When Aesthetics Enhance Perceived Healthiness

Abstract

Marketers frequently style food to look pretty (e.g., in advertising). We investigate how pretty aesthetics (defined by classical aesthetic principles, such as order, symmetry, and balance) influence healthiness judgments. We propose that prettier food is perceived as healthier, specifically because classical aesthetic features make it appear more natural. In a pilot, six main studies, and four supplemental studies (total N = 4,301), across unhealthy and healthy, processed and unprocessed, and photographed and real foods alike, people judged prettier versions of the same food as healthier (e.g., more nutrients, less fat), despite equal perceived price. Even given financial stakes, people were misled by prettiness. Supporting the proposed naturalness process, perceived naturalness mediated the effect; belief in a natural=healthy connection moderated it; expressive aesthetics, which do not evoke naturalness, did not produce the effect (despite being pretty); and reminders of artificial modification, which suppress perceived naturalness, mitigated it. Given that pretty food styling can harm consumers by misleading healthiness judgments for unhealthy foods, managers and policy-makers should consider modification disclaimers as a tool to mitigate the pretty=healthy bias. (176 words)

Keywords: aesthetics; food & health decision-making; inferences; lay intuitions; public policy

People view an estimated 4,013 food and 2,844 restaurant advertisements per year—more than 18 food-related advertisements per day (Statista 2016). In these ads, they typically see foods that are extensively styled to look (unrealistically) pretty. Just imagine the beautiful pizza, with its picture-perfect bubbly crust, flawlessly allocated pepperoni, artfully scattered mushrooms, and glistening pillows of melted cheese, that you see every day driving by one of many billboards. The advertiser's hope is, of course, that prettier foods will seem more appetizing and abundant. Indeed, neuroscience research suggests that viewing delectable food images activates the gustatory cortex, essentially simulating the food's pleasurable taste (Simmons, Martin, and Barsalou 2005). But might pretty aesthetics have other, less obvious, and potentially perverse effects on food evaluations? May the alluringly good-looking pizza actually seem healthier to you, by virtue of its aesthetics? This research examines whether and how prettiness influences judgments of the food's *healthiness*.

Most people strive to eat healthily (International Food Information Council Foundation 2015), but it is notoriously difficult to accurately judge food's healthiness, because nutrients and energy are invisible. Accordingly, consumers' healthiness judgments are often shaped by arbitrary external factors, such as category membership (Rozin, Markwith, and Ashmore 1996), healthy descriptors (Irmak, Vallen, and Robinson 2011), and labels (Schuldt 2013). No research has examined how the prettiness of food impacts healthiness judgments, perhaps because its direction and mechanism are not obvious. On the one hand, beautiful aesthetics may elicit notions of pleasure and hedonics. Activation of these concepts may lead people to view a food as more indulgent, causing them to judge prettier looking foods as unhealthier. On the other hand, beautiful aesthetics may elicit spontaneous inferences of naturalness which, in turn, may signal healthiness. We propose the latter, grounding this proposition in the link between classical

aesthetic principles and patterns found in nature.

Across six experiments that use various different manipulations of prettiness, we find that consumers perceive prettier foods as more natural than less pretty foods, which causes them to infer that prettier foods are healthier. In line with the naturalness hypothesis, people exhibit strong implicit associations between prettiness and naturalness, and naturalness and healthiness. Further, expressive aesthetics, which highlight human intent, do not produce the effect (because they do not evoke naturalness), and disclaimers reminding people that food has been artificially styled mitigate the effect (because they suppress naturalness). Across studies, prettiness only boosts inferences about select attributes, suggesting its effect on perceived healthiness is not due to a general halo. Prettiness also enhances perceived healthiness regardless of perceived tastiness, suggesting the effect cannot be attributed to "wishful thinking." Finally, other inferences that may co-vary with prettiness, such as freshness, care, and sophistication, do not better explain the effect, pointing towards a distinct role of naturalness.

The findings have clear practical relevance. Evaluations of food's healthiness have important downstream consequences, including food choice (Nikolova and Inman 2015) and serving sizes decisions (Suher, Raghunathan, and Hoyer 2016). Given the prominence of food visuals in today's world, paired with marketers' ability to maximize food prettiness in myriad digital and non-digital ways, understanding how aesthetics guide eating decisions is key to promoting healthier choices. For instance, the results point to unnaturalness cues (e.g., reminders about processing) as an intervention to attenuate prettiness-based overestimation of healthiness. Further, the finding that pretty aesthetics can boost perceived healthiness while enhancing (or preserving) perceived tastiness at the same time suggests a way to sidestep the perceived health–taste tradeoff that may deter people from healthy eating (Raghunathan, Naylor, and Hoyer 2006).

The insights also make important theoretical contributions. First, the findings contribute to the literature on lay beliefs by revealing a novel lay intuition: people take classical aesthetics as a signal of naturalness (e.g., being pure, unaltered). Beyond documenting this pretty=natural intuition, they also extend the finding that people believe natural entities to be categorically healthy (Rozin et al. 2004) with objective measures of healthiness (e.g., calories, fat, nutrients).

Second, this research addresses the call for more process research in consumer aesthetics (Patrick and Peracchio 2010): an investigation of different drivers of the effect of pretty aesthetics on judgments offers consistent evidence for the role of naturalness. This finding suggests that classical aesthetics may lead to a host of unintended consequences known to spring from naturalness judgments, and opens up novel avenues for aesthetics research, such as examining the impact of aesthetics on perceived efficacy. Finally, it has implications for any work using visual food stimuli, like that on food choice, health communication, or vice-virtue conflict.

Relationship Between Food Aesthetics and Healthiness Judgments

We define *perceived healthiness* as a composite of being perceived as high in nutrients, low in fat, and low in calories, in addition to being viewed as good for your body and globally healthy, in line with current dietary guidelines and lay understanding of healthiness (International Food Information Council Foundation 2015; United States Department of Agriculture 2015). Because food is complex and the qualities that contribute to its nutritional value are intangible, people rely heavily on a variety of contextual cues to guide their healthiness judgments. For example, mere membership in stereotypically healthy categories enhances perceived healthiness (Rozin et al. 1996). Similarly, healthy-sounding food names (Irmak et al. 2011) and healthy branding of restaurants (Chandon and Wansink 2007) boost the perceived

healthiness of food offerings.

In terms of aesthetic factors and healthiness perceptions, existing research has almost exclusively studied isolated aspects, such as the color of labels (Schuldt 2013) or the shine of packaging (Ye, Morrin, and Kampfer 2020). Whether and how *holistic* food aesthetics (i.e., an overall beautiful, pleasing appearance) shape healthiness perceptions is unknown—a surprising gap, given the prevalence of food styling. Research on the holistic aesthetics of food is scarce in general, and the studies that have examined overall prettiness focus on its effect on food *attractiveness*. For example, in one experiment, people liked a food's taste better when it was visually more appealing (Zellner et al. 2014). In another, diners expressed greater liking and willingness-to-pay prior to eating when a meal was plated more artistically (Michel et al. 2015). In the field, too, consumers tend to avoid ugly-looking food (e.g., De Hooge et al. 2017), which has been attributed to ugly food undermining consumers' self-esteem (Grewal et al. 2019).

This research seeks to fill this gap. We specifically focus on prettiness in terms of *classical* aesthetics—one of two distinct dimensions of human aesthetic perception (see Lavie and Tractinsky 2004). Classical aesthetics are defined by traditional principles of beauty, such as order, symmetry, balance, clarity, and pattern repetition (see also Johnson 1994). We suggest that classical aesthetics signal naturalness, because they resemble patterns found in nature, and thereby enhance perceived healthiness. In contrast, *expressive* aesthetics are marked by the originality and creativity of a designer (Lavie and Tractinsky 2004), and thus are not expected to exert the same effect. Practitioners undoubtedly use aesthetics to make food appear more palatable, but prior research generates multiple possible predictions regarding the unintended and counterintuitive consequences pretty aesthetics may have for perceived healthiness.

Common Assumption: Prettiness Embodies Pleasure and Indulgence

On the one hand, people may intuit that prettier (vs. less pretty) food must be *unhealthy* because prettiness is associated with hedonic pleasure. Beautiful sights, by definition, elicit pleasure. Seeing beautiful designs directly activates the brain's reward center (Reimann et al. 2010), and a primary function of visual art is aesthetic gratification (Hagtvedt and Patrick 2008). The notion of pleasure that is inherent in beautiful visual aesthetics may signal indulgence and, in turn, unhealthiness. People tend to view hedonic attributes as incompatible with instrumental or moral attributes. This principle is reflected in research on self-control dilemmas and vice-virtue conflicts, which readily pits pleasurable options against sensible ones (Okada 2005; Shiv and Fedorikhin 1999).

Some research suggests that the idea of a pleasure–merit tradeoff may be especially salient in the domain of food: people appear to believe that unhealthy food provides greater pleasure, assuming that it is more satisfying (Finkelstein and Fishbach 2010; Suher et al. 2016) and tastier (Raghunathan et al. 2006). Thus, because pretty images are, by definition, pleasurable, pretty (vs. less pretty) versions of the same food may be perceived as *less healthy*.

Proposition: Prettiness Reflects Natural Patterns and Thereby Signals Healthiness

On the other hand, people may perceive prettier (vs. less pretty) food as *healthier*, given that prettier aesthetics often prompt more favorable evaluations on other, unrelated dimensions. For instance, prettier machines, such as ATMs, seem more usable (Tractinsky et al. 2000; for a review see Hassenzahl and Monk 2010); prettier financial reports can elicit higher stock valuation (Townsend and Shu 2010); and even for mundane, machine-made disposable products, like napkins, people believe prettier versions result from greater effort (Wu et al. 2017).

An intuitive explanation for this pattern is a generalized halo effect (Nisbett and Wilson

1977; Thorndike 1920)—a process whereby an initial global positive affective reaction (elicited, e.g., by prettiness) spills over onto the judgment of other, entirely unrelated attributes. Applying this logic here, prettier (vs. less pretty) food may be evaluated more positively on *all* attributes, simply because people feel more positively about it on account of its pleasant looks. However, we suggest that pretty aesthetics provoke distinct inferences about *specific* attributes of the food, which then give rise to perceived healthiness. Specifically, we propose *naturalness* may be a key inference. Diverse literatures hint that prettiness may elicit perceptions of naturalness, and naturalness, in turn, likely boosts perceptions of healthiness, yet this possibility has not been tested.

We propose that prettiness in the form of *classical* aesthetics (i.e., order, symmetry, balance, pattern repetition) induces the notion of naturalness (i.e., being natural, pure, and unprocessed: Rozin 2005), which in turn signals healthiness. Various disciplines have found that patterns that appear in nature are considered beautiful (for a review see Palmer, Schloss, and Sammartino 2013). For instance, approximate symmetry is both extremely prevalent in nature—be it the bilateral symmetry in the vast majority of modern animals (up to 99%: Finnerty 2005) or the radial, bilateral or disymmetry of most flowers (> 90%: Neal, Dafni, and Giurfa 1998)—and considered beautiful (e.g., Jacobson and Höfel 2002; Zheng et al. 2009). Likewise, order and regularity are highly characteristic of nature, emerging in the form of patterns of self-similarity seen in trees, ferns, and other plants as well as coast- and mountain-profiles (see Mandelbrot 1982), and in the pattern repetition exhibited by most visually communicating animals (see Kenward et al. 2004). These, too, are considered beautiful (Spehar et al. 2003). Finally, balanced proportions between different elements or dimensions are common in nature. A prominent example are Fibonacci proportions, which give rise to the "golden spiral" found in most plants' leaf arrangement (Mitchison 1977) and are perceived as beautiful (e.g., Fechner 1871; Green

1995). All of these phenomena are interrelated: pattern repetition produces some order and often gives rise to symmetry; proportionality can induce self-similarity (as it does in logarithmic spirals); and symmetry generates some balance (reflected in "symmetry" meaning "well-proportioned, well-ordered").

In sum, it seems viable that classical aesthetics may elicit the concept of naturalness, and accordingly, we propose that prettier (vs. less pretty) food may be perceived as healthier because it seems more natural. Being natural or unprocessed does not necessarily imply anything about the fat, sugar, or calorie content of a food, but consumers appear to overextend the term to healthiness (Rozin et al. 2004).

Finally, note that the nature-like features of classical aesthetics are not the only way for objects to look pretty. *Expressive* aesthetics (i.e., creativity, originality), for example, refer to the abstract depiction of some imaginative idea—both decidedly human capacities—as opposed to features found in nature. Cubism or abstract expressionism are popular examples of expressive aesthetics. In line with our theorizing that prettiness enhances perceived healthiness by way of naturalness, not all prettiness should produce the effect. We propose that the effect of prettiness on perceived healthiness is elicited by classical aesthetics, which we suggest signal naturalness, but not by other forms (e.g., expressive aesthetics) that do not signal naturalness—despite being pretty.

Other Possible Effects of Prettiness

In addition to the proposed account, the prettiness of a food may inevitably bring to mind any number of other attributes beyond naturalness, and some of these attributes may also be tied to healthiness. For example, perhaps pretty food seems fresher or safer, and people may overgeneralize these qualities to healthiness in terms of calories and nutrients. Pretty food

presentation may also signal food expertise or general refinement. Sophistication is linked with the idea of "upper class" (Aaker 1997), which people may associate with a healthier lifestyle, given that higher socioeconomic status correlates with lower body mass index in developed economies (McLaren 2007); mass media portray success as tantamount to being slim and fit (Evans 2003); and consumers expect expensive foods to be healthier (Haws, Reczek, and Sample 2017). Relatedly, to the extent that pretty food presentation suggests diligent preparation, people may believe that a more caring agent made it. People stereotype others who choose healthy food as more conscientious (e.g., responsible, disciplined, caring: Oakes and Slotterback 2004) and may conversely assume conscientious food presentation is associated with particularly judicious, and healthy, ingredient choices. Neatness is also associated with following prescriptive norms (Vohs, Redden, and Rahinel 2013), which may likewise signal sensible food choices.

Across studies, we examine these possibilities as well as more general alternative mechanisms like a generalized halo effect and motivated reasoning, but none appear to make compelling alternative explanations. We find robust evidence for a distinct effect of naturalness.

Similarly, as we study the effect of visual aesthetics of complex objects (i.e., food) as opposed to isolated elements (e.g., color swatches, dot patterns), the prettier (vs. less pretty) images will typically vary in more than one visual aspect. We have taken several precautions to alleviate a concern that the effect is idiosyncratic to a particular pair of images.

First, a pilot study employs stimulus sampling using images not selected by us. Second, two studies use non-visual manipulations. In one study, all participants evaluate the identical photograph and we manipulate perceived prettiness only by way of expectations to see pretty or ugly food, exploiting an assimilation effect. Another documents the key associations in an Implicit Association Test using only verbal target stimuli. Third, the studies that do manipulate

prettiness visually use various different images (or live food) and yet other stimuli in supplemental studies reported in the Web Appendix. Finally, we measure healthiness as nutrient, fat, and calorie content—variables that should be unaffected by many of these possible differences in the images.

Pilot Study: Suggestive Evidence of the Pretty=Healthy Link

A pilot study (see Web Appendix A) explored if prettier food is indeed perceived as healthier. Eight hundred and three participants were asked to search online for an image of a "pretty" or an "ugly" version of a food (ice cream sundae, burger, pizza, sandwiches, lasagna, omelet, or salad), uploaded the image, and judged to what extent the food pictured was healthy, nutritious, good for me, fatty, and high in calories. The latter two items were reverse-coded and all five items were combined into a "perceived healthiness" scale ($\alpha = .89$).

As predicted, participants perceived the food as healthier when it was pretty ($M = 3.74$, $SD = 1.60$) than when it was ugly ($M = 2.71$, $SD = 1.19$; $F(1, 787) = 209.41$, $p < .001$, $d = .73$; see Table 1; see Web Appendix A for means by food). Gender did not interact with prettiness on any key variable in this or any other study and is thus not discussed further.

Insert Table 1

Of course, the images could have varied in many ways, but this finding provides initial impetus for subsequent experimentation, suggesting a link between prettiness and healthiness with naturalistic stimuli. Study 1A moves to a controlled setting to test the idea that pretty (vs. ugly) food is perceived as healthier.

Study 1A: Prettiness Enhances Perceived Naturalness and Perceived Healthiness

Study 1A manipulates the prettiness of an otherwise identical food by varying the presence of pattern repetition in its presentation but holds constant the setting and photographic quality. Further, to ensure everyone is equally aware of all food components, the stimuli only contain one salient ingredient, and we list ingredients and display the food name with the picture.

Method

Four hundred Prolific Academic panelists in the United States (54.0% women; M_{age} = 35.46, $range_{age}$ = 18–77) were randomly assigned to evaluate either a pretty or an ugly version of an avocado toast (made from the identical ingredients; see Figure 1).

Insert Figure 1

As theories of classical aesthetics predict, in a pretest (see Web Appendix B), people had found the avocado toast significantly prettier when it featured pattern repetition and order (M = 5.24, SD = 1.45) than when it lacked pattern and thus presented no salient order (M = 3.39, SD = 1.58; $p < .001$).

Participants learned that the researchers were interested in opinions about food products and that they would view a food image and rate it on various dimensions. Before viewing the food, participants read that they would see an "avocado toast, made from 1 slice of wheat bread and 1/2 of an avocado." The ingredient text was bolded, and to ensure participants did not skip over these important details, we required them to stay on the page for 10 seconds.

Additionally, on each of the subsequent pages, "Avocado Toast" was displayed under the image alongside a cost of "~$2." They provided all subsequent ratings on 7-point scales anchored at 1 = not at all and 7 = very much.

Perceived healthiness. First, participants rated the avocado toast on the "perceived

healthiness" scale (α = .79; note that this and all subsequent studies use "low in fat" and "low in calories" instead of the reverse-coded items used in the pilot). Then we measured, in randomized order, perceived naturalness and perceived tastiness.

Proposed mediator: Perceived naturalness. To capture the proposed driver of the effect, perceived naturalness, participants rated to what extent they thought the avocado toast was natural, pure, and unprocessed. The items were combined into a "perceived naturalness" scale (α = .87). (Note that principal component analysis showed that naturalness and healthiness are statistically distinct constructs, suggesting discriminant validity; see Web Appendix C.)

Equality check for perceived tastiness. Food styling is done expressly to make the food look pretty and appetizing. Research has shown that consumers believe that unhealthy food is tastier (Raghunathan et al. 2006), so if they also believe the reverse, then styled, pretty food may appear less healthy based merely on this lay theory. For the ideal test, tastiness should be equal across conditions. Participants rated to what extent they thought the avocado toast was tasty, flavorful, and delicious. These items were combined into a "perceived tastiness" scale (α = .97).

Equality check for perceived price. Although equal cost information was provided, to rule out any effects of prettiness on perceived healthiness being due to differences in perceived price (Haws et al. 2017), we measured price perceptions. Participants rated to what extent they thought the avocado toast was pricey and expensive. These items were combined into a "perceived price" composite (r = .92, p < .001).

Lastly, participants reported demographics, which included education level and household income. Neither interacted with prettiness on perceived healthiness or naturalness, so these variables are not discussed further (although details and results are available upon request).

Attention check. Finally, participants responded to an open-ended question asking what

the most prominent component of the food was, to ensure all participants were equally aware that the toast topping was avocado. In this free recall, 98.25% (395 out of 400) answered "avocado;" the five non-avocado responses were distributed equally across conditions ($\chi_2(1) = .21$, $p = .645$).

Results

The means for all dependent variables collected are displayed in Table 1.

Basic effect: Perceived healthiness. Participants rated the avocado toast as significantly healthier when it was pretty (M = 5.03, SD = .98) than when it was ugly (M = 4.70, SD = 1.07; F(1, 398) = 10.04, $p = .002$, d = .32).

Proposed mediator: Perceived naturalness. Participants rated the avocado toast as significantly more natural when it was pretty (M = 5.39, SD = 1.12) than when it was ugly (M = 4.95, SD = 1.37; F(1, 398) = 12.33, $p < .001$, d = .35).

Equality checks for perceived tastiness and price. As intended, people perceived the two avocado toasts as virtually equal in terms of tastiness (F(1, 398) = 2.07, $p = .151$) and perceived price (F(1, 398) = .524, $p = .470$).

Mediation. Simple mediation (model 4; 10,000 samples) indicated that perceived naturalness mediated the effect of prettiness on perceived healthiness (B = .084, SE = .026; 95% CI [.036, .137]; robust to controlling for tastiness and price).

Discussion and Replication Study

People thought the same portion of food was healthier when it was pretty than when it was ugly (despite being given a list of exact ingredients), and this effect was driven by perceived naturalness. Tastiness, in contrast, was unaffected by prettiness.

Replication. We also conducted a conceptual replication (see Web Appendix D) to ensure the generality of the effect and test several alternative explanations. Manipulating

prettiness via order and symmetry of an almond butter and banana toast, it replicated the naturalness-mediated effect of prettiness on perceived healthiness as well as on calorie estimates but found *no* effect of prettiness on perceived tastiness, freshness, amount, or price. Dieting did not moderate the effect.

Together, study 1A and its replication rule out tastiness, freshness, and amount as alternative explanations. It is undoubtedly possible that any of these variables may, for certain stimuli, exert an *additional* positive effect on perceived healthiness beyond that of perceived naturalness, but the results attest that they do not supersede, nor are they necessary for, the mediating effect of naturalness. Beside ruling out these specific alternative mechanisms, the lack of an effect of prettiness on other important attributes (i.e., tastiness, freshness) also challenges a general halo effect as a generic alternative explanation. To demonstrate robustness and eliminate the possibility that any real differences between pretty and ugly images drive the pretty=healthy relationship, study 1B tests the effect using a non-visual manipulation of prettiness.

Study 1B: Prettiness Enhances Perceived Naturalness and Perceived Healthiness

In the pilot study and study 1A, participants evaluated pretty and ugly images that actually differed from each other, which may raise concerns about other attributes that may have co-varied with prettiness. To circumvent this issue, study 1B uses the identical image in both the pretty and the ugly condition and manipulates perceived prettiness of the same image solely by exploiting a biased assimilation effect, whereby the expectation of encountering a certain type of stimulus biases the subsequent perception of the stimulus in the direction of the expectation (Sherif, Taub, and Hovland 1953; for a review of biased assimilation see Lord and Taylor 2009). This type of manipulation eliminates all other objective differences between the foods and varies

prettiness in the cleanest fashion.

Study 1B also includes foods at three different general healthiness levels. The proposed naturalness account does not make direct predictions about any moderating effect of baseline (un)healthiness. Yet, it is possible that very unhealthy categories pose a boundary condition. For example, unhealthy food categories may involve more obviously manipulated components (e.g., processed meats, compounded ingredients like sauces), limiting the extent to which the foods can be construed as natural. Understanding the scope of the effect and its potential for harm are practically important to assess the need for regulation.

Method

Eight hundred and one Prolific Academic panelists in the United States (49.90% women; M_{age} = 32.70, $range_{age}$ = 18–72) were randomly assigned to evaluate food images that they either expected to be pretty or ugly and then, having established this expectation, rated three food replicates, creating a 2(prettiness: pretty vs. ugly)×3(food replicate: almond butter and banana toast, spaghetti marinara, cupcake) mixed design. We included three different foods to test whether the effect of prettiness varies systematically across general healthiness level, and we use a within-subjects design and large sample size to offset the rather subtle manipulation.

Participants learned that the researchers were interested in opinions about food products and that they would view three food images and rate them on various dimensions. To hold the actual food image constant and only change perceived prettiness of the exact same image, we manipulated participants' expectations about whether they would see pretty aesthetic features or ugly aesthetic features by providing the following instructions at the beginning of the study:

> "This study is about PRETTY [UGLY] FOOD.
> The food in the image we will be showing you will be very pretty [ugly] (based on ratings from previous MTurk participants and food design professionals).
> The foods will be orderly [disorderly], they will look symmetrical [lopsided], and the proportions will be balanced [unbalanced]."

In a pretest (see Web Appendix E), people had perceived each food to be equally tasty (p = .791) and equally pricey (p = .901) across the two prettiness conditions.

Then all participants evaluated identical images of an almond butter and banana toast, a plate of spaghetti marinara, and a cupcake, presented in randomized order (see Figure 1). Because moderately ambiguous stimuli are most amenable to manipulation via expectations (see Herr 1989; Sumer and Knight 1996), we selected images that had been rated near the midpoint of the prettiness scale in pretests. Again, the food's name and cost were shown under each image.

For each food, participants first rated the "perceived healthiness" scale ($\alpha_{AB\&B}$ = .83; $\alpha_{Spaghetti}$ = .86; $\alpha_{Cupcake}$ = .91) and then the "perceived naturalness" scale (αs = .89/.90/.87). (Note that principal component analyses showed that naturalness and healthiness are statistically distinct constructs, suggesting discriminant validity; see Web Appendix E.)

Manipulation check: Perceived prettiness. To assess the effectiveness of the manipulation, participants rated to what extent they thought the food was beautiful, pretty, and good looking. The items were combined into a "perceived prettiness" scale (αs = .93/.93/.94). Finally, participants completed an attention check and reported demographics.

Results

The means for all dependent variables collected are displayed in Table 1. We conducted repeated-measures ANOVAs with prettiness as the between-subjects factor and food replicate as the within-subject factor. Prettiness did not interact with food replicate on any of the dependent variables (prettiness: p = .14; healthiness: p = .38; naturalness: p = .37), so we collapsed across the three replicates (but see Web Appendix E for results by food).

Manipulation check: Perceived prettiness. As theories of classical aesthetics and biased assimilation predict, participants found identical food images significantly prettier when they expected to see orderly, symmetrical, and balanced food presentation (M = 4.19, SD = 1.20) than when they expected to see disorderly, lopsided, and unbalanced food presentation (M = 3.66, SD = 1.18; F(1, 799) = 40.29, $p < .001$, d = .45).

Basic effect: Perceived healthiness. Participants rated the identical food images as significantly healthier when they perceived them as pretty (M = 3.56, SD = .73) than when they perceived them as ugly (M = 3.39, SD = .69; F(1, 799) = 11.96, $p = .001$, d = .24).

Proposed mediator: Perceived naturalness. Participants rated the identical food images as significantly more natural when they perceived them as pretty (M = 3.60, SD = .92) than when they perceived them as ugly (M = 3.35, SD = .85; F(1, 799) = 16.38, $p < .001$, d = .28).

Mediation. Because the independent variable of interest, prettiness, was manipulated between-subjects and the lack of an interaction between prettiness and food replicate allowed us to collapse across replicates, we conducted regular mediation. Simple mediation (model 4; 10,000 samples) indicated that perceived naturalness mediated the effect of prettiness on perceived healthiness (B = .07, SE = .017; 95% CI [.033, .100]). Results hold when conducting a separate mediation for each replicate.

Discussion

Replicating the results from study 1A, people judged identical foods as healthier when they looked pretty (vs. ugly) to them, because they perceived them as more natural. The effect of prettiness on perceived naturalness and healthiness was not moderated by the food's general healthiness level. That is, the bias extends to those foods where it is the most problematic.

Importantly, the effect was produced purely by focusing people on order, symmetry, and

balance (vs. disorder, asymmetry, and imbalance) to manipulate perceived prettiness, without any changes in the food picture itself—an especially conservative manipulation. This finding suggests that the effect of prettiness on perceived healthiness does *not* rely on true differences between the food images, but rather on people perceiving the food as featuring classical aesthetics, isolating the psychological underpinnings of the effect beyond any image differences.

While this non-visual manipulation of prettiness enables us to test both the key effect and the proposed mechanism in the cleanest manner, it is also a fairly weak prettiness manipulation (manipulation check d = .45). To test the pretty=natural=healthy framework more effectively, in subsequent studies we return to manipulating prettiness via real visual differences in classical aesthetics between images while holding constant or controlling for other relevant food features.

Studies 2 and 3 explore behavioral consequences of the pretty=healthy effect.

Study 2: Prettiness Increases Willingness-to-Pay for Real Food via Healthiness

Study 2 investigates the effect of prettiness-induced perceived healthiness on a consequential behavior: willingness-to pay (WTP). Studies 1A–B and the replication held constant price, as people may infer healthiness from market price (Haws et al. 2017). Yet, it is possible that people *value* foods they perceive as healthier more highly, irrespective of price (see Sevilla and Townsend 2016 for the price–value distinction in the aesthetics domain). The goal of this study is to quantify to what extent the shift in healthiness judgments contributes to shifts in WTP, *not* to demonstrate the general phenomenon that prettiness enhances WTP (see, e.g., Grewal et al. 2019). Study 2 manipulates prettiness via symmetry and balance of a whole food. It also uses the most unequivocally healthy food (i.e., produce) and ensures that the effect is robust when people interact with real food instead of photographs.

Method

We recruited volunteers for a product evaluation study at a large university on the West Coast of the U.S. in conjunction with "move-in day" over a period of three days (9 hours) for $1 in cash. Eighty-nine student and non-student passersby (57.3% women, M_{age} = 23.61, $range_{age}$ = 17–52) were randomly assigned to evaluate either a pretty or an ugly bell pepper (see Figure 1).

Participants completed the study behind a trifold privacy screen, where a white tray with the ugly or pretty produce had been placed according to condition. They were asked to visually examine the product without touching it. Then they rated, in randomized order, an abbreviated "perceived healthiness" scale (α = .91; only using healthy, nutritious, and good for me, to manage time constraints) and the "perceived tastiness" scale (α = .93).

Willingness-to-pay. Then we elicited WTP using a version of the BDM procedure (Becker, DeGroot, and Marschak 1964). Participants bid some amount of their $1 on the pepper in front of them, anticipating that a die roll would determine its selling price and that if they bid more or equal to the selling price, they would buy the pepper and receive the rest of their $1 in change, but if they bid below the selling price, they would not buy it and instead receive the whole $1 in cash. They chose a bid between $0 and $1 (in 10¢ increments) and displayed this reservation price to the researcher, who then rolled a ten-sided die to determine the selling price.

Manipulation check: Prettiness. Before receiving their cash and/or (a different) pepper, participants rated the "perceived prettiness" scale (α = .97).

Finally, they reported demographics, including income. Income had no main effect and did not interact with prettiness on perceived healthiness or naturalness, so this variable is not discussed further (although details and results are available upon request).

Results

The means (or medians) for all dependent variables collected are displayed in Table 1.

Manipulation check for perceived prettiness. As theories of classical aesthetics predict, participants rated the pepper as significantly prettier when it was symmetrical and balanced (M = 5.41, SD = 1.31) than when it was asymmetrical and unbalanced (M = 3.20, SD = 1.86; F(1, 88) = 43.49, $p < .001$, d = 1.37).

Basic effect: Perceived healthiness. Participants rated the pepper as significantly healthier when it was pretty (M = 6.03, SD = 1.02) than when it was ugly (M = 5.08, SD = 1.49; F(1, 88) = 12.96, $p = .001$, d = .74).

Perceived tastiness. Participants rated the pepper as tastier when it was pretty (M = 5.44, SD = 1.22) than when it was ugly (M = 4.26, SD = 1.68; F(1, 88) = 14.64, $p < .001$, d = .80).

WTP. WTP in the ugly condition was right-skewed (Shapiro-Wilk W (40) = .869, $p < .001$; for the distribution, see Web Appendix F), so we applied a Mann-Whitney U test. People bid significantly more real money to buy the pepper when it was pretty (Md = $.50; M = $.47, SD = .27) than when it was ugly (Md = $.20; M = $.30, SD = .29; U = 1365, z = 3.20, $p < .001$). Results remain unchanged when applying a parametric test (F(1, 88) = 8.96, $p = .004$, d = .61).

Mediation. We tested several mediation models. Separate simple mediations (model 4; 10,000 samples) indicated that perceived healthiness mediated the effect of prettiness on WTP (B = .022, SE = .012; 95% CI [.004, .051]), as did perceived tastiness (B = .031, SE = .014; 95% CI [.009, .063]). Serial mediation with both healthiness and tastiness revealed that a healthiness–tastiness sequence mediated the effect on WTP (B = .012, SE = .008; 95% CI [.001, .030]), but the reversed tastiness–healthiness sequence did *not* (B = .005, SE = .067; 95% CI [–.009, .063]).

Discussion and Replication Study

Once more, people thought food was healthier when it was pretty (vs. ugly), and this

inference raised their willingness to pay real money for the item. Although WTP is undoubtedly multiply determined, perceived healthiness contributed significantly to the increase in valuation.

Replication. We also conducted a conceptual replication online with photographs of peppers and apples (see Web Appendix G). This replication also measured perceived naturalness and found that a naturalness–healthiness sequence serially mediated the effect of prettiness on WTP. It also replicated the finding that healthiness may give rise to tastiness: a naturalness–healthiness–tastiness sequences serially mediated the effect of prettiness on WTP, but reversed sequences (i.e., naturalness–tastiness–healthiness and tastiness–naturalness–healthiness) did *not*.

Study 2 and its replication show that prettiness boosts WTP by increasing perceived healthiness. To demonstrate the effect with a different type of consequential behavior, study 3 tests if the prettiness-induced differences in perceived healthiness are powerful enough to shift people's choices even if they are motivated to choose healthy.

Study 3: Prettiness Impacts Healthiness Judgments Even When Accuracy is Incentivized

Study 2 showed that greater perceived healthiness boosts WTP, but an arguably equally important downstream consequence is choice. People who try to identify healthy options may be misled by unhealthy options (believing them to be the healthy choice), simply by virtue of those unhealthy options looking pretty. Study 3 tests if prettiness biases choice even when people are explicitly incentivized to make a healthy selection.

Method

Three hundred Prolific Academic panel members (53.3% women; M_{age} = 30.80, $range_{age}$ = 18–78) were asked to identify which of two foods had fewer calories. They were randomly assigned to a choice set that contained either a pretty version or an ugly version of a target food

in addition to a reference food, creating a one-factor (target food prettiness: pretty vs. ugly) design. The task was consequential, with a bonus given for the correct choice.

Participants learned that the researchers were interested in how people think about nutrition, that they would be asked to identify the lower calorie food of two options, and that correct answers would be rewarded with a $.25 bonus (doubling their base compensation). Then they saw two foods and chose which one they thought had fewer calories.

Choice. Participants chose between an almond butter and banana (AB&B) toast and an avocado toast, presented in random order. The avocado toast (reference food) was the same for all participants, but for the AB&B toast (target food) they were randomly assigned to see either a pretty version or an ugly version of it (made from the identical ingredients; see Figure 1).

As theories of classical aesthetics predict, in a pretest (see Web Appendix B), people had found the toast significantly prettier when it was orderly and symmetrical (M = 4.14, SD = 1.68) than when it was disorderly and asymmetrical (M = 3.22, SD = 1.74; p = .008).

Again, each food's name was displayed under the respective image along with equal cost information. Participants chose either the AB&B toast (coded as 1) or the avocado toast (coded as 0) as the lower calorie food and reported demographics.

The objectively accurate answer was the avocado toast (which contained 270 calories, whereas the AB&B contained 380 calories). The relevant test is whether people's propensity to (mistakenly) identify the AB&B toast as the lower calorie option is greater when it looks pretty compared to when it looks ugly.

Results

The choice shares for the dependent variable are displayed in Table 1. A logistic regression with target food prettiness as the predictor (ugly serving as the reference category)

and choice of which food contained fewer calories as the dependent variable (the AB&B toast serving as the target criterion) was significant (Wald $\chi_2(1) = 6.63$, B = .613, SE = .238, $p = .010$). Specifically, significantly more people (falsely) identified the AB&B toast as the lower calorie food when they saw its pretty version (48.0% [72 of 150]) than when they saw its ugly version (33.3% [50 of 150]). The odds of choosing the AB&B toast increased by a factor of 1.846 for the pretty version compared to the ugly version (signified by Exp[B]). That is, people were more likely to get it wrong and misidentify the higher calorie food as the lower calorie food merely based on it looking prettier.

Discussion

These results extend the support for our basic hypothesis. In line with pretty foods being judged as healthier in the previous studies, in a consequential task, people were more likely to miscategorize a food as a lower calorie option when it looked pretty than when it looked less pretty. The fact that this effect occurred even when people were financially incentivized to judge accurately suggests it is not due to demand, low effort thinking, or motivated reasoning (which is attenuated by accuracy motivation, see Hart et al. 2009). The phenomenon is robust—people cannot ignore prettiness, even when monetary rewards are at stake.

Cognitive Nature of the Documented Associations and Implications for Moderators

Moving from demonstrating the effect of prettiness on perceived healthiness in various ways, the subsequent studies shift focus to moderators. To make predictions about potential moderation, it is important to first understand the cognitive nature of the associations. For instance, if the effect is rooted in strong implicit associations, interventions targeting the explicit level (e.g., verbal information) will likely require hefty information and deliberation to override

said implicit associations. Therefore, we conducted a supplemental study to explore how the pretty=natural and natural=healthy associations operate psychologically on the explicit and implicit levels (see Web Appendix H).

A combination of Implicit Association Tests (IATs) and surveys revealed that both the pretty=natural and the natural=healthy association are grounded in strong implicit associations. Notably, both associations appear to be much stronger at the implicit level than the explicit level (based on effect sizes). Indeed, for the pretty=natural association, people do not seem to hold a salient explicit belief in either direction. These findings have implications for viable moderators.

First, even people with weak or absent explicit beliefs exhibited implicit associations. Accordingly, while belief strength may moderate the mediating effect of perceived naturalness, weak beliefs alone are unlikely to turn off the effect completely. Indeed, a supplemental moderation study, which manipulated prettiness via order and symmetry using spaghetti marinara, corroborated this hypothesis: a weaker natural=healthy belief weakened the mediation by naturalness, but did not eliminate the pretty=healthy effect (see Web Appendix I).

Second, the best shot at mitigating the robust effect of prettiness on perceived healthiness may therefore lie in turning off the naturalness inference directly. This approach is particularly promising due to the absence of any strong explicit belief about the pretty=natural relationship. One way to curtail naturalness inferences is manipulating prettiness in a manner that does not elicit the notion of naturalness in the first place. Our theory proposes that expressive aesthetics will achieve just that. Another may be reminding people that pretty food has been artificially modified. Studies 4A and 4B investigate these theoretically predicted moderators and boundaries.

Study 4A: Only Classical Aesthetics Enhance Perceived Naturalness and Healthiness

We theorize that classical aesthetics, which resemble patterns found in nature, boost perceived healthiness specifically because they elicit the notion of naturalness. Per this logic, the effect of prettiness on naturalness should not extend to expressive aesthetics, which lack classical aesthetics features and are instead characterized by creativity. Study 4A tests this theoretical prediction. The predicted pattern would also negate perceived sophistication and care as alternate drivers, to the extent they are heightened for expressive aesthetics, and contradict a general halo, as a halo should arise from the generic positive affect from *any* prettiness, not just specific types.

Method

Six hundred and one Amazon Turk panel members in the United States (49.1% women; M_{age} = 36.46, $range_{age}$ = 19–87) were randomly assigned to evaluate one of three photographs: an almond butter and banana (AB&B) toast that was ugly; the same AB&B toast that was pretty by virtue of classic aesthetic features; or an AB&B toast that was pretty by virtue of expressive aesthetic features (made from the same ingredients; see Figure 2).

Insert Figure 2

In a pretest (see Web Appendix B), people had found the AB&B toast significantly prettier when it was symmetrical and orderly (classical aesthetics; M = 4.69, SD = 1.69) and when it displayed basic shapes assembled into a playful scene (expressive aesthetics; M = 4.65, SD = 1.70) than when it was disorderly and asymmetrical (ugly; M = 3.33, SD = 1.76; *p*s < .001), but classical aesthetics and expressive aesthetics were rated as equally pretty (p = .81).

Importantly, the AB&B toast was rated as higher on classical aesthetics in the pretty–classical condition (M = 5.87, SD = 1.04) than in the pretty–expressive condition (M = 4.05, SD = 1.50; $p < .001$) and than in the ugly condition (M = 3.80, SD = 1.58; $p < .001$). Conversely, the

AB&B toast was rated higher on expressive aesthetics in the pretty–expressive condition (M = 5.43, SD = 1.28) than in the pretty–classical condition (M = 4.29, SD = 1.58; $p < .001$) and than in the ugly condition (M = 3.14, SD = 1.66; $p < .001$). As such, the images meet the requirements for testing the hypothesis that the pretty=healthy effect emerges only for prettiness based on classical aesthetics, because they signal naturalness, but not for equally high prettiness based on expressive aesthetics, because they do not.

Before viewing the food, participants read that they would see an "almond butter and banana toast, made from one slice of wheat bread, a scoop of almond butter, and half of a banana." Again, equal cost information was shown under each image.

Then they completed, in randomized order, the "perceived healthiness" scale ($\alpha = .87$), the "perceived naturalness" scale ($\alpha = .85$), and measures of perceived sophistication and care. However, neither the results for sophistication nor those for care aligned with the results for perceived healthiness, so measures and detailed results are only reported in Web Appendix J (but see Table 1 for means). Finally, participants reported demographics.

Results

The means for all dependent variables collected are displayed in Table 1.

Basic effect: Perceived healthiness. Condition had a significant effect ($F(2, 598) = 10.82$, $p < .001$). Compared to the ugly condition (M = 4.83, SD = 1.13), people rated the toast significantly healthier in the pretty–classical condition (M = 5.31, SD = 1.09; $t(598) = –4.24$, $p < .001$, d = .43), but *not* any healthier in the pretty–expressive condition (M = 4.88, SD = 1.17; $t(598) = –.44$, $p = .663$, d = .04). They also rated the toast as significantly healthier in the pretty—classical than in the pretty—expressive condition ($t(598) = –3.79$, $p < .001$, d = .38).

Proposed mediator: Perceived naturalness. Condition had a significant effect (F(2, 598)

= 14.53, $p < .001$). Compared to the ugly condition (M = 4.90, SD = 1.27), people rated the toast significantly more natural in the pretty–classical condition (M = 5.45, SD = 1.19; t(598) = –4.23, $p < .001$, d = .45), but *no* more natural in the pretty–expressive condition (M = 4.80, SD = 1.39; t(598) = .80, $p = .425$, d = .08). They also rated the toast as more natural in the pretty–classical than in the pretty–expressive condition (t(598) = –5.02, $p < .001$, d = .50).

Mediation. Simple mediation (multi-categorical independent variable with indicator coding, the ugly condition serving as the reference category; model 4; 10,000 samples) with perceived naturalness as the mediator returned the following results. For the comparison between the ugly and the pretty–classical aesthetics conditions, naturalness mediated the effect of prettiness on perceived healthiness (B = .292, SE = .068; 95% CI [.164, .428]). For the comparison between the ugly and the pretty–expressive conditions, naturalness did *not* mediate (B = –.055, SE = .070; 95% CI [–.194, .082])—as expected, given that these groups differed neither on naturalness nor healthiness. An additional mediation with pretty–expressive serving as the reference category showed that for the comparison between the pretty–expressive aesthetics and the pretty–classical aesthetics conditions, naturalness also mediated the effect of prettiness on perceived healthiness (B = .347, SE = .072; 95% CI [.205, .487]).

The patterns of the perceived sophistication and perceived care results across the three conditions did not align with the perceived healthiness results, making them unlikely mediators. Nonetheless, we also tested parallel mediation as above with naturalness, care, and sophistication simultaneously. Central to testing the theory about the role of perceived naturalness, all mediation patterns for naturalness remained unchanged, that is, naturalness mediated for all three comparisons above and beyond the other variables. Care also had some additional explanatory value for some comparisons; sophistication had none. Because the key insight is that naturalness

is a *unique* driver of the effect, these ancillary results are reported only in Web Appendix J.

Discussion

These results not only provide more support for the proposition that prettiness in the form of classical aesthetics boosts perceived healthiness by signaling naturalness, but they also show that neither sophistication nor care are viable alternative explanation. Further, both the classical and expressive pretty toast being rated as equally pretty also speaks against general positive inferences from general attractiveness as an alternative explanation and contradicts the idea of prettiness merely casting a general halo.

Study 4B: Artificial Modification Disclaimer Eliminates the Effect of Prettiness

We propose that prettiness enhances perceived healthiness by way of intuitively being perceived as more natural. Accordingly, reminders that the food was artificially styled should mitigate the effect of prettiness on perceived healthiness, insofar as they suppress naturalness perceptions. Study 4B tests this theoretical prediction. The anticipated pattern would provide further evidence of process by moderation and also point to an effective intervention. Study 4A manipulated (classical) aesthetics chiefly via symmetry and order. Study 4B manipulates prettiness via the presence of pattern repetition, while again listing ingredients and displaying the food name with the picture.

Method

Three hundred and one Prolific Academic panelists in the United States (55.8% women; M_{age} = 34.17, $range_{age}$ = 18–79) were randomly assigned to evaluate one of three photos: an ugly avocado toast, a pretty avocado toast, or the pretty avocado toast plus a disclaimer that the food had been artificially modified (all made from the identical ingredients; see Figure 2).

In a pretest (see Web Appendix B), people had found the avocado toast significantly prettier both when it featured pattern repetition (M = 4.57, SD = 1.57) and when it featured pattern repetition plus the disclaimer (M = 4.41, SD = 1.56) than when it lacked salient pattern (M = 3.06, SD = 1.46; *p*s < .001), but found the two pretty conditions equally pretty (*p* = .591).

Before viewing the food, all participants read that they would see an "avocado toast, made from 1 slice of wheat bread and 1/2 an avocado." The ingredient text was bolded, and to ensure participants did not skip over these important details, participants were required to stay on the page for 10 seconds. Participants in the pretty+disclaimer condition additionally read that the food they would see had been "artificially modified for advertising," and "as a result, the food may strike [them] as unnatural." Then participants moved on to viewing the food.

Again, the food's name and equal cost information were shown under each image. Participants in the pretty+disclaimer condition additionally saw a statement re-iterating that the food was "artificially modified for advertising" and that it was "not a natural representation" (see Figure 2). The Federal Trade Commission requires that advertising disclaimers be "clear and conspicuous" (ftc.gov). Thus, presenting the statement near the relevant content resembles how it would likely be displayed in real life.

They completed the "perceived healthiness" scale (α = .83), and then, in randomized order, the "perceived naturalness" scale (α = .86) and the "perceived tastiness" scale (α = .95). They also completed the "perceived price" measure (*r* = .87, *p* < .001). As in studies 1A–B and the replication, tastiness and price were equal across conditions (see Web Appendix K).

They also rated a measure of perceived amount. However, the pattern did not align with that for perceived healthiness, so measures and detailed results are only reported in Web Appendix K (but see Table 1 for means).

Lastly, participants reported demographics, which included education level and household income. Neither interacted with prettiness on perceived healthiness or naturalness, so these variables are not discussed further (although details and results are available upon request).

Attention check. Finally, participants recalled the most prominent component of the food as in study 1A. In this free recall, 98.67% (297 out of 301) responded "avocado;" the four other responses were distributed equally across conditions ($\chi^2(2) = .51$, $p = .773$).

Results

The means for all dependent variables collected are displayed in Table 1.

Basic effect: Perceived healthiness. Condition had a significant effect ($F(2, 298) = 4.57$, $p = .011$). Compared to the ugly condition (M = 4.58, SD = 1.22), people rated the avocado toast as significantly healthier in the pretty condition (M = 5.07, SD = 1.13; $t(298) = –2.99$, $p = .003$, d = .42), but *not* any healthier in the pretty+disclaimer condition (M = 4.76, SD = 1.16; $t(298) = –1.11$, $p = .267$, d = .15). They also rated the toast marginally healthier in the pretty condition than in the pretty+disclaimer condition ($t(298) = –1.87$, $p = .062$, d = .27).

Proposed mediator: Perceived naturalness. Condition had a significant effect ($F(2, 298) = 8.78$, $p < .001$). Compared to the ugly condition (M = 4.89, SD = 1.27), people rated the toast as significantly more natural in the pretty condition (M = 5.52, SD = 1.04; $t(298) = –3.38$, $p = .001$, d = .54), but *no* more natural in the pretty+disclaimer condition (M = 4.81, SD = 1.60; $t(298) = –.45$, $p = .656$, d = .06). They also rated the toast as significantly more natural in the pretty condition than the pretty+disclaimer condition ($t(298) = –3.83$, $p < .001$, d = .53).

Mediation. Simple mediation (multi-categorical independent variable with indicator coding, ugly serving as the reference category; model 4; 10,000 samples) with perceived naturalness as the mediator returned the following results. For the comparison between the pretty

and the ugly conditions, naturalness mediated the effect of prettiness on perceived healthiness (B = .285, SE = .080; 95% CI [.136, .447]). For the comparison between the pretty+disclaimer and the ugly conditions, naturalness did *not* mediate (B = –.038, SE = .093; 95% CI [–.221, .147])—as expected, given that the groups differed on neither naturalness nor healthiness. An additional mediation with pretty+disclaimer serving as the reference category showed that for the comparison between the pretty and the pretty+disclaimer conditions, naturalness also mediated the effect of prettiness on perceived healthiness (B = .323, SE = .089; 95% CI [.155, .504]). All results are robust to controlling for tastiness, amount, and price.

Discussion

These results are informative from both a theoretical and a managerial or policy perspective. First, they provide process-by-moderation evidence for the idea that prettiness drives up perceived healthiness specifically by signaling naturalness: as this theory predicts, if the pretty=natural link is disrupted, prettiness no longer exerts its effect. Similar to study 1B, this pattern also rules out several alternative explanations, such as actual differences in food visibility or amounts or prototypicality of the food presentation between the visuals, given that the two pretty conditions used exactly the same photograph, yet still produced healthiness differences based on the artificial modification disclaimer. Beyond the moderation pattern, tastiness was unaffected by prettiness, challenging both a motivated reasoning and a general halo explanation.

Second, these results offer important directions to marketers or policy-makers who want to protect consumers from drawing false inferences about healthiness. Evidently, consumers can be inoculated against the misleading effect of aesthetics, specifically by highlighting that food in advertising pictures has been artificially modified. Disclaimers placed near food images may be an actionable way to mitigate the elusory healthiness boost that prettiness otherwise induces.

General Discussion

People constantly encounter food that is styled to look pretty. The literature makes multiple compelling, but divergent, predictions: pretty aesthetics may elicit notions of pleasure and thereby induce *lower* healthiness judgments, or pretty aesthetics may give rise to intuitions about specific attributes that lead to *higher* healthiness judgments. We propose the latter, and suggest that perceived naturalness may be a key driver. Six high-powered experiments (N = 2,492), a pilot study (N = 803), and four supplemental studies (N = 1,006) demonstrate that people perceive prettier (vs. less pretty) versions of the same food as healthier because they seem more natural. This effect materializes with naturalistic and controlled stimuli; with visual and non-visual manipulations of prettiness; with photographed and live, unhealthy and healthy, and processed and whole foods; it misleads people's choices even when they have financial stakes; and emerges irrespective of prettier food looking tastier and larger. Prettiness only affects select attributes, and only classical aesthetics produce the effect, contradicting a general halo. The effect is independent of tastiness and the persists when accuracy is incentivized, speaking against motivated reasoning. In all studies, the effect emerges despite equal price perceptions.

Theoretical Contribution

This research is the first to explore the role of aesthetics in healthiness judgments. We systematically tested different possible processes that could give rise to the observed effect of aesthetics, and found the effect of aesthetics on perceived healthiness has a cognitive basis (lay intuitions), rather than an affective (halo effect) or motivational ("wishful thinking") one.

These findings add to a body of work that documents how lay theories of food, nutrition, taste, and dieting govern many everyday food decisions (Finkelstein and Fishbach 2010; Haws et al. 2017; McFerran and Mukhopadhyay 2013; Raghunathan et al. 2006). First and foremost, they

reveal the novel lay intuition that classical aesthetics are a sign of naturalness. They also support and extend work on the basic link between naturalness and healthiness, which has shown that more natural entities are ranked as generally "healthier" than less natural entities (Rozin et al. 2004), by quantifying health perceptions in a more nuanced way (e.g., calorie and fat content). Further, this research identifies a moderator of the unhealthy=tasty belief. Prior work, in which foods were either not depicted or visually identical, found that describing foods in unhealthy (vs. healthy) terms may enhance expected and experienced tastiness (Raghunathan et al. 2006; but see Mai and Hoffmann 2015). Yet, our research suggests that classical aesthetics can make food seem simultaneously healthier *and* tastier. This qualification is critical, as it offers a solution to overcome the (perceived) health–taste tradeoff that may pose a barrier to healthy choices.

This research shows that classical aesthetics elicit a sense of naturalness. This association is somewhat counterintuitive—achieving the maxims of classical aesthetics (i.e., order, symmetry, balance) often requires effort and artificial manipulation. Yet, it raises novel implications for aesthetics research. A new way to think about "what is beautiful is good"-type effects may rather be that "what is *natural* is good:" naturalness may underlie other positive aesthetics effects attributed to a "beauty halo." The results also highlight that aesthetics research may benefit from distinguishing between classical and expressive aesthetics, as they may operate differently. We offer initial evidence that the pretty=natural effect may be limited to classical aesthetics, but additional empirical examination of the different types of aesthetics is warranted.

The pretty=natural effect in particular opens up new avenues for aesthetics research. Not only may the effect generalize from food to other entities (e.g., products, people), it also points to novel related mediators (e.g., authenticity, talent) and downstream effects of aesthetics in other domains. For example, it is conceivable that interior design emphasizing classical aesthetic

principles may cast products sold or services rendered in a space as more wholesome and health-promoting. It also prompts questions about how aesthetics impact perceived product efficacy. Some research suggests that outcomes achieved naturally are perceived as higher quality (e.g., Tsay 2016), but other work hints that more natural products are perceived as weaker (e.g., Luchs et al. 2010). How aesthetics, by way of influencing perceived naturalness, shape consumers' beliefs about product success or strength should be an important line of inquiry for marketers.

Marketing and Policy Implications

The fact that the same food appears healthier when it looks more aesthetically pleasing than when it looks less aesthetically pleasing has practical implications for marketers and policy-makers. First, it means that many food advertisements and restaurant menus may, by virtue of depicting heavily styled foods, be promising more than tastiness alone. They may be (falsely) heralding greater levels of healthiness along with it. Our research showed that the pretty=healthy effect extends to unhealthy foods (e.g., pastries). This finding is disconcerting because a large proportion of visually advertised food is unhealthy food. For example, about 72% of restaurant ads viewed per year are fast food restaurants', and the most advertised food brand in the U.S. is McDonald's (Statista 2016). If fast food is consistently portrayed in ways that increase perceived healthiness, even consumers who are motivated to select foods based on healthfulness may become more likely to make unhealthy choices, such as considering fast food as an option at all, choosing it more frequently, selecting larger portions of seemingly not-so-unhealthy foods, or under-estimating how much "balancing out" fast food requires (e.g., via exercise). Transcending unhealthy choices at individual decision points, the persistent subtle boost in perceived healthiness may, over time, even promote overly optimistic general beliefs about unhealthy food types, for instance, that fast food overall is fairly healthy.

Second, food advertisers and restaurant marketers may strategically leverage aesthetics as a signal of healthiness in times when health-related marketing language is increasingly under scrutiny. Since 2016, the Food and Drug Administration has begun curtailing which foods can be labeled "healthy," but healthfulness is a valuable attribute (with losses as much as $3.82 million in monthly revenue after dropping health claims: Rao and Wang 2017). Firms may seek more surreptitious ways to elevate perceived healthiness among their increasingly health-conscious customer base. The potential use of aesthetics as a (deceptive) signal that misleads "reasonable" persons (i.e., college students, campus passersby, online panelists) warrants close consideration by policy-makers. Even companies that do not wish to actively deceive customers may inadvertently lead people astray in their healthiness judgments, simply by presenting food in a (classically) pretty fashion. Indeed, in this research, we found that consumers do *not* seem to hold a salient belief that prettier food is more natural, and this lack of awareness may make them particularly vulnerable (and marketing managers oblivious) to the bias.

On the upside, our research also identifies an effective intervention to protect consumers from being deceived by foods' aesthetics. We found that a statement that explicitly reminds people that a pretty food was artificially modified for depiction can mitigate the effect. Given that it is not viable to prohibit firms from depicting food products in a basic aesthetically pleasant manner altogether, disclaimers may be the most practical solution.

Limitations and Future Research

This research was designed to test the influence of food aesthetics on perceived healthiness and food choice. Food evaluation and choice are critical for a healthy lifestyle, but of course, subsequent food consumption decisions, such as portion size selection or intake, play a role as well. We did not examine actual consumption, and given the complexity of eating

decisions we expect multiple judgments to interact with each other. For instance, people may eat less of a high (vs. low) aesthetic food for fear of destroying its beauty (Wu et al. 2017, study 2) in spite of its apparent healthiness. At the same time, people often overcompensate and increase consumption when food seems healthy (Suher et al. 2016) or when context factors reduce guilt for unhealthy eating (Hagen, Krishna, and McFerran 2017), so they may eat more of prettier food because it seems healthier. Future research may investigate these competing forces directly.

Our findings also cannot fully illuminate whether the effect of prettiness is uneven in the very low versus the very high prettiness range. We found the effect with stimuli at various levels of prettiness (see Web Appendix B), which suggests the effect is not limited to either the positive or the negative realm. Nonetheless, it is conceivable that prettiness has a stronger effect in the below-average domain (i.e., very ugly vs. medium) than the above-average domain (i.e., medium vs. very pretty), akin to patterns found in person perception research (Griffin and Langlois 2006).

Relatedly, our studies demonstrate the effect of prettiness on perceived healthiness across a variety of foods, but there are likely boundary conditions. For example, the effect emerged with unhealthy food categories (e.g., pizza, cupcakes), but may not extend to blatantly unhealthy food (e.g., pretty butter pats); and it appeared with highly processed foods (e.g., frosting), but may well be thwarted by extremely unnatural components (e.g., neon dyes, edible glitter).

Finally, this research documents the intuitions that people have, but not their origin—are they merely overgeneralizations from otherwise true correlations in the real world or utter misconceptions about the food system? It is unrealistic that prettier food really is more natural on average, given how much preparation and additional ingredients go into making food look pretty (e.g., color preservatives). Instead, an extrapolation from nature-like patterns that define classical aesthetics to naturalness seems reasonable. Likewise, greater naturalness does not necessarily

boost nutritional value (Smith-Spangler et al. 2012), but it is plausible that consumers frequently observe a co-occurrence of naturalness signals (e.g., organic label) and healthiness signals (e.g., marketers' emphasis of a healthy lifestyle). Content analysis and survey data may be a more fruitful approach to the complex issue of how lay beliefs develop than experimental methods.

Conclusion

This research exposes a novel effect of aesthetics in the domain of food and reveals an unrecognized influence on healthiness judgments. People perceive the same food as more natural when it happens to look prettier and believe this naturalness implies healthiness—both the presence of positive elements (e.g., nutrients) and the absence of negative elements (e.g., calories). The investigation used primarily laboratory experiments and focused chiefly on *causes* of consumers' perceptions of healthiness rather than on the *consequences* of these perceptions. However, the studies document that perceived healthiness impacts real willingness-to-pay and that the pretty=healthy bias is robust even when consumers are (financially) motivated to choose the healthy option. While the origins of lay intuitions about nutrition are not always clear, other research has shown that they can meaningfully influence outcomes ranging from food choice to consumption to weight status (McFerran and Mukhopadhyay 2013). Given the prevalent use of food visuals in marketing practice and research alike, understanding the effect of food aesthetics offers important insights into the thought processes of consumers trying to make pretty healthy choices.

References

Aaker, Jennifer L. (1997), "Dimensions of Brand Personality," *Journal of Marketing Research*, 34 (3), 347–56.

Becker, Gordon M., Morris H. DeGroot, and Jacob Marschak (1964), "Measuring Utility by a Single-Response Sequential Method," *Behavioral Science*, 9 (3), 226–32.

Carpenter, Thomas P., Ruth Pogacar, Chris Pullig, Michal Kouril, Stephen Aguilar, Jordan LaBouff, Naomi Isenberg, and Alek Chakroff (2019), "Survey-Software Implicit Association Tests: A Methodological and Empirical Analysis," *Behavior Research Methods*, 51 (5), 2194–2208.

Chandon, Pierre, and Brian Wansink (2007), "The Biasing Health Halos of Fast-food Restaurant Health Claims: Lower Calorie Estimates and Higher Side-dish Consumption Intention," *Journal of Consumer Research*, 34 (3), 301–14.

De Hooge, Ilona E., Marije Oostindjer, Jessica Aschemann-Witzel, Anne Normann, Simone Mueller Loose, and Valérie Lengard Almli (2017), "This Apple is Too Ugly for Me!: Consumer Preferences for Suboptimal Food Products in the Supermarket and at Home," *Food Quality and Preference*, 56, 80–92.

Fechner, Gustav T. (1871), *Zur Experimentalen Aesthetik*, Leipzig: Hirzel.

Finkelstein, Stacey R., and Ayelet Fishbach (2010), "When Healthy Food Makes You Hungry," *Journal of Consumer Research*, 37 (3), 357–67.

Finnerty, John R. (2005), "Did Internal Transport, Rather than Directed Locomotion, Favor the Evolution of Bilateral Symmetry in Animals?," *BioEssays*, 27 (11), 1174–180.

Green, Christopher D. (1995), "All that Glitters: A Review of Psychological Research on the Aesthetics of the Golden Section," *Perception*, 24 (8), 937–68.

Greenwald, Anthony G., Debbie E. McGhee, and Jordan L. K. Schwartz (1998), "Measuring Individual Differences in Implicit Cognition: The Implicit Association Test," *Journal of Personality and Social Psychology*, 74 (June), 1464–80.

Greenwald, Anthony G., Brian A. Nosek, and Mahzarin R. Banaji (2003), "Understanding and

Using the Implicit Association Test: I. An Improved Scoring Algorithm," *Journal of Personality and Social Psychology*, 85 (2), 197–216.

Grewal, Lauren, Jillian Hmurovic, Cait Lamberton, and Rebecca Walker Reczek (2019), "The Self-Perception Connection: Why Consumers Devalue Unattractive Produce," *Journal of Marketing*, 83 (1), 89–107.

Griffin, Angela M., and Judith H. Langlois (2006), "Stereotype Directionality and Attractiveness Stereotyping: Is Beauty Good or Is Ugly Bad?," *Social Cognition*, 24 (2), 187–206.

Hagen, Linda, Aradhna Krishna, and Brent McFerran (2017), "Rejecting Responsibility: Low Physical Involvement in Obtaining Food Promotes Unhealthy Eating," *Journal of Marketing Research*, 54 (4), 589–604.

Hagtvedt, Henrik, and Vanessa M. Patrick (2008), "Art Infusion: The Influence of Visual Art on the Perception and Evaluation of Consumer Products," *Journal of Marketing Research*, 45 (3), 379–89.

Hart, William, Dolores Albarracín, Alice H. Eagly, Inge Brechan, Matthew J. Lindberg, and Lisa Merrill (2009), "Feeling Validated versus Being Correct: A Meta-analysis of Selective Exposure to Information," *Psychological Bulletin*, 135 (4), 555–88.

Hassenzahl, Marc, and Andrew Monk (2010), "The Inference of Perceived Usability from Beauty," *Human-Computer Interaction*, 25 (3), 235–60.

Haws, Kelly L., Rebecca Walker Reczek, and Kevin L. Sample (2017), "Healthy Diets Make Empty Wallets: The Healthy = Expensive Intuition," *Journal of Consumer Research*, 43 (6), 992–1007.

Herr, Paul M. (1989), "Priming Price: Prior Knowledge and Context Effects," *Journal of Consumer Research*, 16 (June), 67-75.

International Food Information Council Foundation (2015), *Food & Health Survey 2015*. Retrieved from https://foodinsight.org

Irmak, Caglar, Beth Vallen, and Stefanie Rosen Robinson (2011), "The Impact of Product Name

on Dieters' and Nondieters' Food Evaluations and Consumption," *Journal of Consumer Research*, 38 (2), 390–405.

Jacobson, Thomas, and Lea A. Höfel (2002), "Aesthetic Judgments of Novel Graphic Patterns: Analyses of Individual Judgments," *Perceptual and Motor Skills*, 95 (3), 755–66.

Johnson, Paul-Alan (1994), *The Theory of Architecture: Concepts, Themes, and Practices*, New York: Van Nostrand Reinhold.

Kenward, Benjamin, Carl-Adam Wachtmeister, Stefano Ghirlanda, and Magnus Enquist (2004), "Spots and Stripes: The Evolution of Repetition in Visual Signal Form," *Journal of Theoretical Biology*, 230 (3), 407-419.

Lavie, Talia, and Noam Tractinsky (2004), "Assessing Dimensions of Perceived Visual Aesthetics of Web Sites," *International Journal of Human-Computer Studies*, 60, 269–98.

Lord, Charles G., and Cheryl A. Taylor (2009), "Biased Assimilation: Effects of Assumptions and Expectations on the Interpretation of New Evidence," *Social and Personality Psychology Compass*, 3 (5), 827–41.

Luchs, Michael G., Rebecca W. Naylor, Julie R. Irwin, and Rajagopal Raghunathan (2010), "The Sustainability Liability: Potential Negative Effects of Ethicality on Product Preference," *Journal of Marketing*, 74 (5), 18–31.

Mai, Robert, and Stefan Hoffmann (2015), "How to Combat the Unhealthy = Tasty Intuition: The Influencing Role of Health Consciousness," *Journal of Public Policy & Marketing*, 34 (1), 63–83.

Mandelbrot, Benoît B. (1982), *The Fractal Geometry of Nature*. New York: W. H. Freeman.

McFerran, Brent, and Anirban Mukhopadhyay (2013), "Lay Theories of Obesity Predict Actual Body Mass," *Psychological Science*, 24 (8), 1428–436.

McLaren, Lindsay (2007), "Socioeconomic Status and Obesity," *Epidemiologic Reviews*, 29 (1), 29–48.

Michel, Charles, Carlos Velasco, Paul Fraemohs, and Charles Spence (2015), "Studying the

Impact of Plating on Ratings of the Food Served in a Naturalistic Dining Context," *Appetite*, 90, 45–50.

Mitchison, Graeme (1977), "Phyllotaxis and the Fibonacci Series," *Science*, 196 (4287), 270–75.

Neal, Paul R., Amots Dafni, and Martin Giurfa (1998), "Floral Symmetry and Its Role in Plant-pollinator Systems: Terminology, Distribution, and Hypotheses," *Annual Review of Ecology and Systematics*, 29 (1), 345–73.

Nikolova, Hristina Dzhogleva, and J. Jeffrey Inman (2015), "Healthy Choice: The Effect of Simplified Point-of-Sale Nutritional Information on Consumer Food Choice Behavior," *Journal of Marketing Research*, 52 (6), 817–35.

Nisbett, Richard E., and Timothy D. Wilson (1977), "The Halo Effect: Evidence for Unconscious Alteration of Judgments," *Journal of Personality and Social Psychology*, 35 (4), 250–56.

Oakes, Michael E., and Carole S. Slotterback (2004), "Prejudgments of Those Who Eat a 'Healthy' versus an 'Unhealthy' Food for Breakfast," *Current Psychology*, 23 (4), 267–77.

Okada, Erica M. (2005), "Justification Effects on Consumer Choice of Hedonic and Utilitarian Goods," *Journal of Marketing Research*, 42 (1), 43–53.

Palmer, Stephen E., Karen B. Schloss, and Jonathan Sammartino (2013), "Visual Aesthetics and Human Preference," *Annual Review of Psychology*, 64, 77–107.

Patrick, Vanessa M., and Laura A. Peracchio (2010), "'Curating' the JCP Special Issue on Aesthetics in Consumer Psychology: An Introduction to the Aesthetics Issue," *Journal of Consumer Psychology*, 20 (4), 393–97.

Raghunathan, Rajagopal, Rebecca Walker Naylor, and Wayne D. Hoyer (2006), "The Unhealthy = Tasty Intuition and Its Effects on Taste Inferences, Enjoyment, and Choice of Food Products," *Journal of Marketing*, 70 (4), 170–84.

Rao, Anita, and Emily Wang (2017), "Demand for "Healthy" Products: False Claims and FTC Regulation," *Journal of Marketing Research*, 44 (6), 968–89.

Reimann, Martin, Judith Zaichkowsky, Carolin Neuhaus, Thomas Bender, and Bernd Weber (2010), "Aesthetic Package Design: A Behavioral, Neural, and Psychological

Investigation," *Journal of Consumer Psychology,* 20 (4), 431–41.

Rozin, Paul (2005), "The Meaning of "Natural": Process More Important than Content," *Psychological Science*, 16 (8), 652–58.

Rozin, Paul, Michele Ashmore, and Maureen Markwith (1996), "Lay American Conceptions of Nutrition: Dose Insensitivity, Categorical Thinking, Contagion, and the Monotonic Mind," *Health Psychology*, 15 (6), 438–47.

Rozin, Paul, Mark Spranca, Zeev Krieger, Ruth Neuhaus, Darlene Surillo, Amy Swerdlin, and Katherine Wood (2004), "Natural Preference: Instrumental and Ideational/Moral Motivations, and the Contrast between Foods and Medicines," *Appetite*, 43 (2), 147–54.

Schuldt, Jonathon P. (2013), "Does Green Mean Healthy? Nutrition Label Color Affects Perceptions of Healthfulness," *Health Communication,* 28 (8), 814–21.

Sevilla, Julio, and Claudia Townsend (2016), "The Space-to-Product Ratio Effect: How Interstitial Space Influences Product Aesthetic Appeal, Store Perceptions, and Product Preference," *Journal of Marketing Research*, 53 (5), 665–81.

Sherif, Muzafer, Daniel Taub, and Carl I. Hovland (1958), "Assimilation and Contrast Effects of Anchoring Stimuli on Judgments," *Journal of Experimental Psychology*, 55 (2), 150-55.

Shiv, Baba, and Alexander Fedorikhin (1999), "Heart and Mind in Conflict: The Interplay of Affect and Cognition in Consumer Decision Making," *Journal of Consumer Research,* 26 (3), 278–92.

Simmons, Kyle, Alex Martin, and Lawrence W. Barsalou (2005), "Pictures of Appetizing Foods Activate Gustatory Cortices for Taste and Reward," *Cerebral Cortex,* 15 (10), 1602–608.

Smith-Spangler, Crystal, et al. (2012), "Are Organic Foods Safer or Healthier Than Conventional Alternatives? A Systematic Review," *Annals of Internal Medicine*, 157 (5), 348–66.

Spehar, Branka, Colin W. G. Clifford, Ben R. Newell, and Richard P. Taylor (2003), "Universal Aesthetic of Fractals," *Computers & Graphics*, 27 (5), 813–20.

Statista (2016), *Food Advertising in the United States*. Retrieved from

https://www.statista.com/study/24278/food-advertising-in-the-us-statista-dossier/

Suher, Jacob, Raj Raghunathan, and Wayne D. Hoyer (2016), "Eating Healthy or Feeling Empty? How the "Healthy = Less Filling" Intuition Influences Satiety," *Journal of the Association for Consumer Research*, 1 (1), 26–40.

Thorndike, Edward L. (1920), "A Constant Error in Psychological Ratings," *Journal of Applied Psychology*, 4 (1), 25-29.

Townsend, Claudia, and Suzanne B. Shu (2010), "When and How Aesthetics Influence Financial Decisions," *Journal of Consumer Psychology*, 20 (4), 452–58.

Tractinsky, Noam, Adi Shoval-Katz, and D. Ikar (2000), "What is Beautiful is Usable," *Interacting with Computers*, 13 (2), 127–45.

Tsay, Chia-Jung (2016), "Privileging Naturals over Strivers: The Costs of the Naturalness Bias," *Personality and Social Psychology Bulletin*, 42 (1), 40–53.

United States Department of Agriculture (2015), *Dietary Guidelines for Americans 2015–2020 Eight Edition*. Retrieved from https://health.gov/dietaryguidelines/2015/

Wu, Freeman, Adriana Samper, Andrea C. Morales, and Gavan J. Fitzsimons (2017), "It's Too Pretty to Use! When and How Enhanced Product Aesthetics Discourage Usage and Lower Consumption Enjoyment," *Journal of Consumer Research*, 44 (3), 651–72.

Ye, Ning, Maureen Morrin, and Kristina Kampfer (2020), "From Glossy to Greasy: The Impact of Learned Associations on Perceptions of Food Healthfulness," *Journal of Consumer Psychology*, 30 (1), 96–124.

Zellner, Debra A., Christopher R. Loss, Jonathan Zearfoss, and Sergio Remolina (2014), "It Tastes as Good as It Looks! The Effect of Food Presentation on Liking for the Flavor of Food," *Appetite*, 77, 31–35.

Zheng, Xianjun Sam, Ishani Chakraborty, James Jeng-Weei Lin, and Robert Rauschenberger (2009), "Correlating Low-level Image Statistics with Users' Rapid Aesthetic and Affective Judgments of Web Pages," in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, 1–10.

TABLE 1
RESULTS OF ALL MAIN TEXT STUDIES

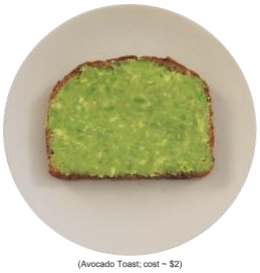
Study	Prettiness Manipulation (Food)	Measure	Condition			Pr > Test Stat
			Ugly	Pretty		
Pilot	Search online for image of pretty (vs. ugly) food (Various; e.g., salad, sandwich, pizza)	Prettiness	1.46 (.77)	6.14 (.94)		$p < .001$
		Healthiness	2.71 (1.19)	3.74 (1.60)		$p < .001$
1A	Pattern repetition (Avocado toast)	Healthiness	4.70 (1.07)	5.03 (.98)		$p = .002$
		Naturalness	4.95 (1.37)	5.39 (1.12)		$p < .001$
		Tastiness	4.29 (1.94)	4.56 (1.88)		$p = .151$
		Price	3.61 (1.83)	3.47 (1.89)		$p = .470$
1B	Expectations/assimilation (Almond butter & banana toast; spaghetti marinara; cupcake)	Prettiness	3.66 (1.18)	4.19 (1.20)		$p < .001$
		Healthiness	3.39 (.69)	3.56 (.73)		$p = .001$
		Naturalness	3.35 (.85)	3.60 (.92)		$p < .001$
2	Symmetry + balance (Bell pepper)	WTP	Md \$.20 \$.30 (.29)	Md \$.50 \$.47 (.27)		$p < .001$ $p = .004$
		Healthiness	5.08 (1.49)	6.03 (1.02)		$p = .001$
		Tastiness	4.26 (1.68)	5.44 (1.22)		$p < .001$
3	Order + symmetry (Almond butter & banana toast)	Choice as lower calorie item	33.3%	48.0%		$p = .010$
			Ugly	Pretty—Classical Aesthetics	Pretty—Expressive Aesthetics	
4A	Order + symmetry (Almond butter & banana toast)	Healthiness	4.83 (1.13)$_a$	5.31 (1.09)$_b$	4.88 (1.17)$_a$	$p < .001$
		Naturalness	4.90 (1.27)$_a$	5.45 (1.19)$_b$	4.80 (1.39)$_a$	$p < .001$
		Sophistication	2.70 (1.25)$_a$	3.40 (1.39)$_b$	4.33 (1.33)$_c$	$p < .001$
		Care	4.72 (1.07)$_a$	5.53 (1.06)$_b$	5.60 (1.07)$_b$	$p < .001$
			Ugly	Pretty	Pretty + Disclaimer	
4B	Pattern repetition (Avocado toast)	Healthiness	4.58 (1.22)$_a$	5.07 (1.13)$_{b,\dagger}$	4.76 (1.16)$_{a,\dagger}$	$p = .011$
		Naturalness	4.89 (1.27)$_a$	5.52 (1.04)$_b$	4.81 (1.60)$_a$	$p < .001$
		Tastiness	3.83 (1.84)$_a$	4.23 (1.86)$_a$	3.92 (1.99)$_a$	$p = .294$
		Amount	3.49 (1.28)$_a$	4.23 (1.42)$_b$	4.27 (1.49)$_b$	$p < .001$
		Price	3.62 (1.78)$_a$	3.28 (1.77)$_a$	3.48 (1.78)$_a$	$p = .410$

Notes: Results are means (SD) unless otherwise indicated. For studies 4A and 4B, means in the same row with different superscripts are significantly different from each other at the $p < .05$ level; means that share the † symbol are significantly different from each other at the $p < .1$ level; means that share the same superscript letter do not differ from each other.

FIGURE 1: STIMULI USED IN STUDIES 1–3

Study 1A

Ugly Condition



(Avocado Toast; cost ~ $2)

Pretty Condition

(Avocado Toast; cost ~ $2)

Note: The avocado toasts are the identical food (one slice of bread and half an avocado). However, the pretty avocado toast (center and right) features clear pattern repetition and, by extension, presents some explicit order, whereas the ugly avocado toast (left) lacks pattern and presents no markers of order.

This study is about UGLY FOOD.
The food in the image we will be showing you will be very ugly (based on ratings from previous MTurk participants and food design professionals).
The foods will be disorderly, they will look lopsided, and the proportions will be unbalanced.

This study is about PRETTY FOOD.
The food in the image we will be showing you will be very pretty (based on ratings from previous MTurk participants and food design professionals).
The foods will be orderly, they will look symmetrical, and the proportions will be balanced.

Study 1B

(Almond butter & banana toast; cost ~ $2)

(Spaghetti marinara; price: ~ $9.50)

(Cupcake; cost ~ $1)

Note: The three food images (presented in random order) were identical in both conditions. Prettiness was manipulated by inducing expectations to see either pretty or ugly food pictures. The pretty condition (right) was led to expect order, symmetry, and balance, whereas the ugly condition (left) was led to expect disorder, lopsidedness, and imbalance. (Note that the cupcake picture is from Wu et al. 2017)

Study 2

Note: Real bell peppers participants bid on. The pretty pepper (right) is symmetrical and evenly shaped, whereas the ugly pepper (left) is asymmetrical and unevenly shaped.

Study 3

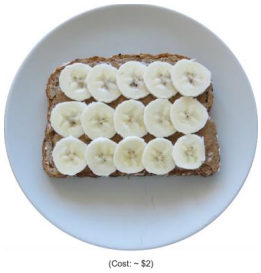
Note: The target food is the almond butter & banana (AB&B) toast; the toasts are the identical food (one slice of bread, one scoop of almond butter, and 15 banana slices). The pretty AB&B toast (right) is orderly and symmetrical with clean, parallel visual lines, whereas the ugly AB&B toast (left) is disorderly and asymmetrical with broken, disparate lines. The avocado toast serves as the reference food and the image is the same in both conditions. Presentation order of the options was randomized.

FIGURE 2: STIMULI USED IN STUDIES 4A AND 4B

Study 4A

Ugly Condition

Pretty–Classical Condition

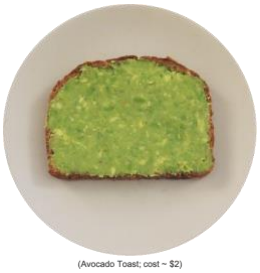


Pretty–Expressive Condition

Note: Almond butter and banana (AB&B) toasts are made from the same ingredients. The pretty—classical aesthetics AB&B toast (center) is arranged in an orderly, symmetrical pattern with clean, parallel visual lines, whereas the pretty—expressive aesthetics AB&B (right) toast is arranged in an imaginative way representing a scene with houses, trees, and stars. The ugly AB&B toast (left) is arranged in a disorderly, asymmetrical way with broken, disparate visual lines.

Study 4B

Ugly Condition



Pretty Condition

Pretty+Disclaimer Condition

Note: Avocado toasts are the identical food (one slice of bread, one scoop of almond butter, and 15 banana slices). However, the pretty avocado toasts (center and right) feature clear pattern repetition and, by extension, presents some explicit order, whereas the ugly avocado toast (left) lacks pattern and presents no markers of order.

Pretty Healthy Food: How and When Aesthetics Enhance Perceived Healthiness

Web Appendix

A. Pilot Study
B. Pretests for Stimuli Used in Studies and Replications
C. Study 1A Supplementary Analyses
D. Study 1A Conceptual Replication: Prettiness via Symmetry and Order Increases Healthiness and Reduces Calorie Estimates via Naturalness
E. Study 1B Pretest and Supplementary Analyses
F. Study 2 Supplementary Analyses
G. Study 2 Conceptual Replication: Prettiness Boosts Willingness-to-Pay via Naturalness and Healthiness
H. Supplemental Study: Cognitive Nature of the Pretty=Natural=Healthy Associations
I. Supplemental Moderation Study: Stronger Mediating Effect Among Those with a Stronger Natural=Healthy Belief
J. Study 4A Supplementary Methods and Analyses
K. Study 4B Supplementary Methods and Analyses

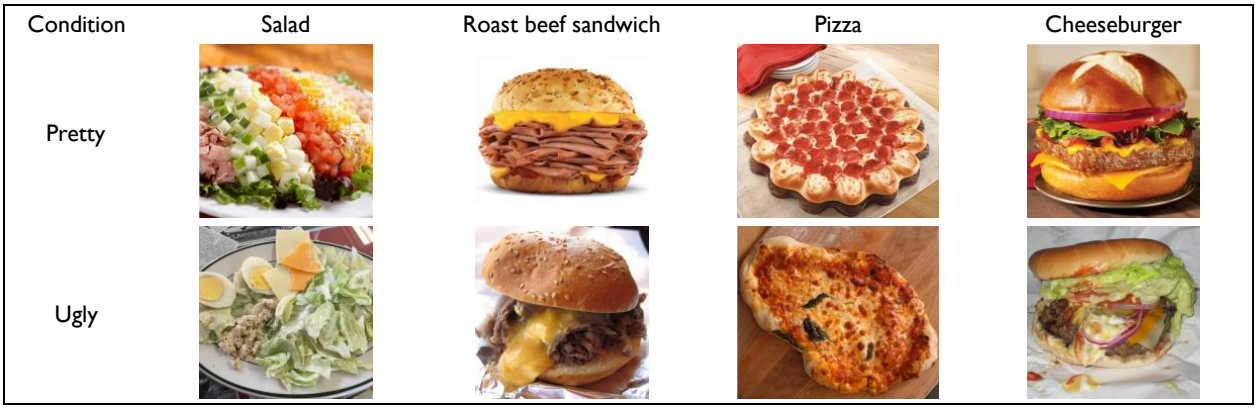
A: PILOT STUDY

The pilot study explored if prettier food is indeed perceived as healthier. For this initial probe, we asked people to search online and evaluate an image of a pretty or an ugly version of a food and measured perceived healthiness.

Method

Eight hundred and three Amazon Turk panel members (41.3% women; M_{age} = 32.71, $range_{age}$ = 18–73) were randomly assigned to search online, upload, and evaluate an image of a pretty or an ugly version of one of eight foods (ice cream sundae, burger, pizza, sandwiches, lasagna, omelet, salad), creating a 2(prettiness: pretty vs. ugly)×8(food type) between-subjects design. "Pretty" ["ugly"] was defined as "aesthetically pleasant [unpleasant] or presented in a good- [bad-]looking way." The image had to show the food itself, and the food had to be real, edible, and focal. This task guards against the concern that effects of prettiness are merely driven by idiosyncratic features of any given picture: in principle, it generates as many different stimuli as there are participants (i.e., 803), which should wash out any effects of any specific detail; it has ecological validity, as people find the association "in the wild;" and we had no hand in selecting the stimuli (see Figure S1 for examples of food pictures that respondents uploaded).

FIGURE S1
EXAMPLES OF FOOD PICTURES UPLOADED BY PARTICIPANTS IN PILOT



Participants then evaluated the food. They provided all subsequent ratings on 7-point scales anchored at 1 = not at all and 7 = very much.

Perceived healthiness. We first measured the focal dependent variable, perceived healthiness. Participants rated to what extent they thought the food pictured was healthy, nutritious, good for me, fatty, and high in calories. The latter two items were reverse-coded and all five items were combined into a "perceived healthiness" scale (α = .89).

Manipulation check for perceived prettiness. Participants rated to what extent they thought the food was beautiful, pretty, and good looking. The items were combined into a "perceived prettiness" scale (α = .99). (Note that principal component analysis showed that prettiness and healthiness are statistically distinct constructs, suggesting discriminant validity; see below.) Finally, participants completed an attention check and reported demographics.

Results and Discussion

Prettiness and gender did not interact on the focal dependent variable, healthiness; thus, gender is not included in the analyses presented here.

Manipulation check for perceived prettiness. As intended, participants rated the food as prettier when asked to select a pretty version (M = 6.14, SD = .94) compared to an ugly version (M = 1.46, SD = .77; F(1, 787) = 6014.09, $p < .001$, d = 5.45; contrast Fs > 574.25, ps < .001). As such, the prettiness manipulation was successful.

Basic effect: Perceived healthiness. As predicted, participants perceived the food as healthier when it was pretty (M = 3.74, SD = 1.60) than when it was ugly (M = 2.71, SD = 1.19; F(1, 787) = 209.41, $p < .001$, d = .73). Given the range of foods, there was also an incidental main effect of food type (F(1, 787) = 86.60, $p < .001$) and a prettiness×food type interaction

(F(1, 787) = 8.75, $p < .001$). For seven of eight foods, participants rated the food as healthier when it was pretty than when it was ugly (Fs > 9.19, ps < .003); for one, the pattern was in the predicted direction but not significant ($F_{icecream} = 1.44$, $p = .231$; see Table S1 for means by food), supporting the idea that prettier (vs. ugly) foods are perceived to be healthier.

TABLE S1
MEAN (SD) PERCEIVED HEALTHINESS BY CONDITION IN PILOT

Condition	Salad	Omelet	Fish sandwich	Lasagna	Roast beef sandwich	Pizza	Cheese-burger	Ice cream sundae
Pretty	6.15 (.75)	4.87 (.98)	4.10 (1.20)	3.56 (1.04)	3.53 (1.35)	3.18 (1.19)	2.60 (1.10)	2.19 (.90)
Ugly	3.95 (1.23)	3.28 (1.08)	2.97 (1.14)	2.80 (.92)	2.20 (.78)	2.47 (1.04)	1.99 (.81)	1.92 (1.00)

Principal component analysis. Principal component analysis indicated that a two-factor solution best represented the eight items: two factors had eigenvalues above 1 (4.54 and 2.17), and the second factor explained a considerable amount of variance (27.06%) beyond what the first factor explained (56.79%). The factor loadings are displayed in Table S2. The "perceived prettiness" items loaded together on one factor; the "perceived healthiness" items on the other.

TABLE S2
FACTOR LOADINGS OF HEALTHINESS AND PRETTINESS ITEMS IN THE PILOT STUDY AFTER VARIMAX ROTATION

Item	Component	
	1	2
healthy	.405	.815
nutritious	.481	.737
good for me	.491	.749
fatty$_R$	-.046	.855
high calorie$_R$	-.158	.844
beautiful	.962	.102
pretty	.965	.112
good looking	.971	.101

B: PRETESTS FOR STIMULI USED IN STUDIES AND REPLICATIONS

Method

All pretests were conducted separately. Participants (see Table S3 for sample descriptives) were randomly assigned to view either the pretty or the ugly food picture used in the respective main study (including the food's name and cost information; see respective studies for images) and rated the following measure(s) on 7-point scales anchored at 1 = not at all and 7 = very much.

Perceived prettiness. Participants in all pretests rated to what extent they thought the food pictured was beautiful, pretty, and good looking (combined into a "perceived prettiness" scale; see Table S3 for Cronbach's alphas by study).

Participants in the pretest for study 4A rated two additional measures of aesthetics.

Classical aesthetics. Beyond this general prettiness evaluation, they rated how symmetrical, orderly, and balanced the image looked to them. These items were combined into a "classical aesthetics" scale ($\alpha = .89$).

Expressive aesthetics. They also rated how imaginative and artistic they found the image. These items were combined into an "expressive aesthetics" composite ($r = .76$, $p < .001$).

Results

All results are displayed in Table S3.

TABLE S3
RESULTS OF PRETTINESS PRETEST RESULTS

Study	Prettiness Manipulation (Food)	Sample Descriptives	Measure	Condition			F-test	T-test Against Scale Midpoint (4)
				Ugly	Pretty			
1A	Pattern repetition (Avocado toast)	111 MTurkers (52.3% women; M_{age} = 35.96, $range_{age}$ = 18–67)	Prettiness α = .97	3.39 (1.58)	5.24 (1.44)		$F(1, 109) = 41.36$, $p < .001$	$t_{ugly}(54) = -2.87$, $p = .006$ $t_{pretty}(55) = 6.41$, $p < .001$
1A Conceptual Replication	Order + symmetry (Almond butter & banana toast)	97 MTurkers (41.2% women; M_{age} = 34.86, $range_{age}$ = 18–77)	Prettiness α = .90	2.77 (1.38)	3.73 (1.40)		$F(1, 95) = 11.51$, $p = .001$	$t_{ugly}(46) = -6.08$, $p < .001$ $t_{pretty}(49) = -1.34$, $p = .19$
2 Conceptual Replication	Symmetry + balance (Bell pepper, apple)	405 MTurkers (40.5% women; M_{age} = 34.65, $range_{age}$ = 19–71)	Pepper Prettiness α = .98	2.69 (1.65)	5.86 (1.25)		$F(1, 202) = 238.13$, $p < .001$	$t_{ugly}(99) = -7.98$, $p < .001$ $t_{pretty}(101) = 15.06$, $p < .001$
			Apple Prettiness α = .97	3.59 (1.89)	6.08 (1.07)		$F(1, 199) = 131.99$, $p < .001$	$t_{ugly}(99) = -2.17$, $p = .032$ $t_{pretty}(100) = 19.47$, $p < .001$
3	Order + symmetry (Almond butter & banana toast)	101 MTurkers (48.5% women; M_{age} = 35.48, $range_{age}$ = 18–70)	Prettiness α = .96	3.22 (1.74)	4.14 (1.68)		$F(1, 99) = 7.28$, $p = .008$	$t_{ugly}(50) = -3.19$, $p = .002$ $t_{pretty}(49) = .59$, $p = .557$
Additional Moderation Study	Order + symmetry (Spaghetti marinara)	100 MTurkers (46.0% women; M_{age} = 34.71, $range_{age}$ = 18–67)	Prettiness α = .97	2.68 (1.67)	4.50 (1.56)		$F(1, 98) = 31.81$, $p < .001$	$t_{ugly}(48) = -5.53$, $p < .001$ $t_{pretty}(50) = 2.30$, $p = .025$
				Ugly	Pretty—Classical Aesthetics	Pretty—Expressive Aesthetics		
4A	Order + symmetry (Almond butter & banana toast)	597 MTurkers (54.6% women; M_{age} = 37.04, $range_{age}$ = 18–75)	Prettiness α = .96	3.33_a (1.76)	4.69_b (1.69)	4.65_b (1.69)	$F(2, 594) = 40.45$, $p < .001$	$t_{ugly}(197) = -5.37$, $p < .001$ $t_{pr\text{-}class}(199) = 5.78$, $p < .001$ $t_{pr\text{-}expr}(198) = 5.40$, $p < .001$
			Class. Aesth. α = .89	3.80_a (1.58)	5.87_b (1.04)	4.05_a (1.50)		
			Express. Aesth. r = .76, p<.001	3.14_a (1.66)	4.29_b (1.58)	5.43_c (1.28)		
				Ugly	Pretty	Pretty + Disclaimer		
4B	Pattern repetition (Avocado toast)	151 Prolific panelists (60.3% women; M_{age} = 35.99, $range_{age}$ = 18–72)	Prettiness α = .93	3.06_a (1.46)	4.57_b (1.57)	4.41_b (1.56)	$F(2, 148) = 14.48$, $p < .001$	$t_{ugly}(48) = -4.50$, $p < .001$ $t_{pretty}(50) = 2.59$, $p = .013$ $t_{pr\text{-}discl}(50) = 1.85$, $p = .070$

Notes: Results are means (SD) unless otherwise indicated. For studies 4A and 4B, means in the same row with different superscripts are significantly different from each other at the $p < .05$ level; means that share the same superscript letter are not different from each other.

C: STUDY 1A SUPPLEMENTARY ANALYSES

Results

Principal component analysis. Principal component analysis indicated that a multi-factor solution best represented the eight items measuring perceived naturalness and perceived healthiness. Three factors had an eigenvalue above 1 (4.10, 1.55, and 1.20), and the second (19.40%) and third (14.94%) factors explained a considerable amount of variance beyond what the first factor explained (51.27%). The factor loadings are displayed in Table S4.

The three "perceived naturalness" items loaded together on one factor, and the positive and negative "perceived healthiness" items loaded on one factor, respectively. Given the good internal consistency of the five healthiness items, we suspect this is an artifact of the negative versus positive framing of the questions rather than their content. (The results of the principal component analyses for study 1B and the replication of study 2, which used the positively framed items "low in fat" and "low in calories" supports this idea; see Web Appendices E and G.) More importantly, the results indicate that naturalness and healthiness are statistically distinct constructs.

TABLE S4
FACTOR LOADINGS OF NATURALNESS AND HEALTHINESS ITEMS IN STUDY 1A AFTER VARIMAX ROTATION

Item	Component		
	1	2	3
nutritious	.912	.232	.073
healthy	.892	.241	.173
good for me	.883	.263	.097
unprocessed	.135	.866	.156
natural	.278	.862	.043
pure	.329	.834	.084
low calorie	.131	.081	.934
low fat	.102	.124	.832

D: STUDY 1A CONCEPTUAL REPLICATION: PRETTINESS INCREASES PERCEIVED HEALTHINESS AND REDUCES CALORIE ESTIMATES VIA NATURALNESS

In study 1A, prettiness was manipulated via the presence (vs. absence) of pattern repetition. If the effect of prettiness on perceived healthiness is indeed rooted in classical aesthetics, it should emerge for other forms of classical aesthetic principles. This replication manipulates the prettiness of an otherwise identical food by varying the amount of visual order and symmetry in its presentation, but again holds constant the setting and photographic quality. Further, the stimuli are equal on perceived price, amount, tastiness, and freshness.

Method

Two hundred and one Prolific Academic panelists in the United States (55.2% women; M_{age} = 35.98, $range_{age}$ = 18–79) were randomly assigned to evaluate an almond butter and banana (AB&B) toast that looked pretty or the identical toast that did not look pretty (see Figure S2), creating a one-factor (prettiness: pretty vs. ugly) design.

FIGURE S2
STIMULI USED IN THE CONCEPTUAL REPLICATION OF STUDY 1A

Note: Almond butter and banana (AB&B) toasts. The pretty AB&B toast (right) is arranged in an orderly, symmetrical pattern with clean, parallel visual lines, whereas the ugly AB&B toast (left) is arranged in a disorderly, asymmetrical way with broken, disparate visual lines.

As theories of classical aesthetics predict, in a pretest (see Web Appendix B), people had found the toast significantly prettier when it featured an orderly, symmetrical pattern (M = 3.73, SD = 1.40) than when it was disorderly and asymmetrical (M = 2.77, SD = 1.38; p = .001). In the pretest, even without price labels, people had expected the toast to be priced virtually equally whether it was pretty (M = 1.91, SD = 1.28) or ugly (M = 1.98, SD = 1.30; p = .763). Still, to ensure equal cost perceptions, a cost of "~$2" was shown under each image.

They received the same instructions as in study 1A and provided the following ratings.

Perceived healthiness and calorie estimate. First, people rated "perceived healthiness" (α = .81) as in the pilot study. Additionally, they provided a numerical calorie estimate in response to the prompt, "Please estimate how many calories you think this food contains."

Proposed mediator: Perceived naturalness. Then, they rated the "perceived naturalness" scale (α = .88) as in study 1A. (Note that principal component analysis showed that naturalness and healthiness are statistically distinct constructs, suggesting discriminant validity; see below.)

In addition to the dependent variables of interest, we ensured both conditions were equal on three variables that might vary with prettiness and that could impact perceived healthiness.

Equality check for perceived tastiness. Participants rated the "perceived tastiness" scale (α = .93) as in study 1A.

Equality check for perceived amount. Food styling also typically involves making the food ingredients look plumper and portions more generous, so the styled, pretty food may appear less healthy merely by virtue of looking larger. For the ideal test, amount should be equal across conditions. Participants rated to what extent they thought the food was filling, satiating, large, and substantial. These items were combined into a "perceived amount" scale (α = .86).

Equality check for perceived freshness. Food styling may make food appear fresher, and

thus safer to eat, so pretty food may be seen as healthier via an overgeneralization from not being noxious to being very nutritive. This idea is distinct from the proposed naturalness account. Participants rated to what extent they thought the food was fresh and stale. The latter was reverse coded and the items were combined into a "perceived freshness" composite ($r = .44$, $p < .001$).

Equality check for perceived price. Although the toasts were seen as equally costly in the pretest, and explicit, equal cost information was given in the main study, we measured "perceived price" ($r = .91$, $p < .001$) as in study 1A.

Dietary concern. To test if dieters—who are motivated to distort evaluations to render food acceptably healthy (e.g., Mohr, Lichtenstein, and Janiszewski 2012)—exhibit a stronger effect, we measured dietary concern. If motivated reasoning is at play, dieters should exhibit a stronger effect, as they are more driven to distort evaluations to render food acceptably healthy (e.g., Mohr, Lichtenstein, and Janiszewski 2012). Participants completed the four-item Dietary Concern scale (Mohr et al. 2012), rating how often they watch the amounts of calories and fat they consume; moderate their sugar intake; and cut back on snacks and treats on 6-point scales anchored at 1 = none of the time and 6 = all of the time ($\alpha = .83$). Finally, participants reported demographics.

Results and Discussion

Basic effect: Perceived healthiness. Participants rated the AB&B toast as significantly healthier when it was pretty (M = 5.01, SD = .99) than when it was ugly (M = 4.43, SD = 1.16; $F(1, 199) = 14.32$, $p < .001$, d = .54).

Basic effect: Calorie estimate. Participants also estimated the AB&B toast to contain significantly fewer calories when it was pretty (M = 249.55, SD = 124.43) than when it was ugly (M = 292.20, SD = 135.87; $F(1, 199) = 5.39$, $p = .021$, d = .33).

Proposed mediator: Perceived naturalness. Participants rated the AB&B toast as significantly more natural when it was pretty (M = 4.90, SD = 1.34) than when it was ugly (M = 4.40, SD = 1.30; F(1, 199) = 7.14, p = .008, d = .38).

Equality checks for perceived tastiness, amount, freshness, and price. As intended, people perceived the two AB&B toasts as virtually equal in terms of tastiness (F(1, 199) = .74, p = .391), amount (F(1, 199) = .25, p = .620), freshness (U = 5661, z = 1.54, p = .123) and price (U = 5509.5, z = 1.24, p = .255; note that the latter two variables were non-normally distributed, so we applied a Mann-Whitney U test, although results remain unchanged when applying parametric tests), regardless of the toast's prettiness.

Dietary concern. Linear regressions with centered prettiness condition, centered dietary concern, and their interaction showed that *only* prettiness had an effect on perceived healthiness (B = .303, SE = .077; t(197) = 3.94, p < .001) and naturalness (B = .253, SE = .094; t(197) = 32.68, p = .008). Dietary concern had no main effect on perceived healthiness or naturalness (ps >.16) and did not interact with prettiness on either (ps > .68).

Mediation. Simple mediation (model 4; 10,000 samples) indicated that perceived naturalness mediated both the effect of prettiness on perceived healthiness (B = .192, SE = .077; 95% CI [.059, .370]; robust to controlling for tastiness, amount, price, and freshness) and the effect of prettiness on calorie estimate (B = –11.063, SE = 5.206; 95% CI [–24.274, –3.117]; robust to controlling for tastiness, amount, price, and freshness).

Replicating the results of the pilot and study 1A with a controlled visual manipulation of prettiness, people thought the same portion of food was healthier, and contained fewer calories, when it was pretty than when it was ugly. Two other important attributes, tastiness and freshness, on the other hand, were unaffected by prettiness, casting doubt on a general halo effect.

Principal component analysis. Principal component analysis indicated that a multi-factor solution best represented these eight items measuring perceived naturalness and perceived healthiness. Three factors had an eigenvalue above 1 (4.15, 1.46, and 1.17), and the second (18.22%) and third (14.66%) factors explained a considerable amount of variance beyond what the first factor explained (51.90%). The factor loadings are displayed in Table S5.

The three "perceived naturalness" items loaded together on one factor, and the positive and negative "perceived healthiness" items loaded on one factor, respectively. Given the good internal consistency of the five healthiness items, we suspect this is an artifact of the negative versus positive framing of the questions rather than their content. (The results of the principal component analyses for study 1B and the replication of study 2, which used the positively framed items "low in fat" and "low in calories" supports this idea; see Web Appendices E and G.) More importantly, the results indicate that naturalness and healthiness are statistically distinct constructs.

TABLE S5
FACTOR LOADINGS OF NATURALNESS AND HEALTHINESS ITEMS IN CONCEPTUAL REPLICATION OF STUDY 1A AFTER VARIMAX ROTATION

Item	Component		
	1	2	3
healthy	.907	.238	.184
good for me	.902	.250	.131
nutritious	.897	.244	.079
unprocessed	.096	.893	.098
natural	.308	.857	.075
pure	.385	.811	.063
fatty$_R$	.096	.057	.903
high calorie$_R$	.158	.106	.882

E: STUDY 1B PRETEST AND SUPPLEMENTARY ANALYSES

Pretest

Method. One hundred ninety-eight Prolific Academic panelists in the United States (52.0% women; M_{age} = 30.60, $range_{age}$ = 18–70) were randomly assigned to evaluate food images that they either expected to be pretty or ugly and then rated three food replicates, creating a 2(prettiness: pretty vs. ugly)×3(food replicate: almond butter and banana [AB&B] toast, spaghetti marinara, cupcake) mixed design.

Equality check for perceived tastiness. Food styling is done expressly to make the food look pretty and appetizing. Research has shown that consumers believe that unhealthy food is tastier (Raghunathan et al. 2006), so if they also believe the reverse, then pretty food may appear less healthy based merely on this lay theory. Participants rated to what extent they thought the food was tasty, flavorful, and delicious. These items were combined into a "perceived tastiness" scale ($\alpha_{AB\&B}$ = .94; $\alpha_{Spaghetti}$ = .94; $\alpha_{Cupcake}$ = .92).

Equality check for perceived price. To rule out any effects of prettiness on perceived healthiness being due to differences in perceived price (Haws et al. 2017), we measured price perceptions. Participants rated to what extent they thought the food was pricey and expensive. These items were combined into a "perceived price" composite ($r_{AB\&B}$ = .93, $p < .001$; $r_{Spaghetti}$ = .92, $p < .001$; $r_{Cupcake}$ = .93, $p < .001$).

Results. We conducted repeated-measures ANOVA with prettiness as the between-subjects factor and food replicate as the within-subject factor. For perceived tastiness, there was only an incidental main effect of food replicate ($F(2, 195) = 5.37$, $p = .005$). There was no effect of prettiness ($F(1, 196) = .071$, $p = .791$) and no prettiness×food replicate interaction ($F(2, 195)$

= .16, p = .855).

For perceived price, there was only an incidental main effect of food replicate (F(2, 195) = 5.37, p = .005). There was no effect of prettiness (F(1, 196) = .015, p = .901) and an incidental prettiness×food replicate interaction (F(2, 195) = 3.48, p = .033). Following up on the interaction revealed that perceived price was not significantly different for any of the food replicates ($p_{AB\&B}$ = .092; $p_{Spaghetti}$ = .933; $p_{Cupcake}$ = .259), but apparently, the different direction of the differences as well as the different sizes of the differences gave rise to an interaction.

Supplementary Results

Means by food. All results are displayed in Table S6.

TABLE S6
MEANS (SD) BY FOOD IN STUDY 1B

Measure	Condition		F-test
	Ugly	Pretty	
Prettiness (overall)	3.66 (1.18)	4.19 (1.20)	F(1, 799) = 40.29, p < .001
AB&B toast	4.04 (1.55)	4.47 (1.53)	F(1, 799) = 15.55, p < .001
Spaghetti	3.57 (1.53)	4.08 (1.45)	F(1, 799) = 22.30, p < .001
Cupcake	3.36 (1.53)	4.02 (1.63)	F(1, 799) = 35.45, p < .001
Healthiness (overall)	3.39 (.69)	3.56 (.73)	(1, 799) = 11.96, p = .001
AB&B toast	4.89 (.99)	5.02 (1.01)	F(1, 799) = 3.09, p = .079
Spaghetti	3.62 (1.12)	3.85 (1.03)	F(1, 799) = 9.55, p = .002
Cupcake	1.66 (.80)	1.82 (.97)	F(1, 799) = 6.83, p = .009
Naturalness (overall)	3.35 (.85)	3.60 (.92)	F(1, 799) = 16.38, p < .001
AB&B toast	4.94 (1.26)	5.13 (1.21)	F(1, 799) = 4.47, p = .035
Spaghetti	3.28 (1.27)	3.60 (1.33)	F(1, 799) = 12.50, p < .001
Cupcake	1.83 (1.03)	2.08 (1.16)	F(1, 799) = 10.87, p = .001

Principal component analyses. Principal component analyses were conducted separately for each stimulus. The factor loadings are displayed in Table S7.

For the spaghetti marinara and the cupcake stimuli, two-factor solutions best represented the eight items. For the spaghetti, two factors had eigenvalues above 1 (4.63 and 1.14), and the second factor explained a considerable amount of variance (14.29%) beyond what the first factor explained (57.99%). For the cupcake, two factors had eigenvalues above 1 (4.99 and 1.10), and

the second factor explained a considerable amount of variance (13.68%) beyond what the first factor explained (62.38%). For both foods, the “perceived naturalness” items loaded together on one factor; the “perceived healthiness” items on the other.

For the AB&B toast, a multi-factor solution best represented these eight items measuring perceived naturalness and perceived healthiness. Three factors had an eigenvalue above 1 (4.45, 1.27, and 1.10), and the second (15.88%) and third (13.62%) factors explained a considerable amount of variance beyond what the first factor explained (55.63%). The three “perceived naturalness” items loaded together on one factor, and the positive and negative “perceived healthiness” items loaded on one factor, respectively. Given the good internal consistency of the five healthiness items, and in light of the two-factor solutions for the other two stimuli, we suspect this is an artifact of the negative versus positive framing of the questions rather than their content. More importantly, the results indicate that naturalness and healthiness are statistically distinct constructs.

TABLE S7
FACTOR LOADINGS OF NATURALNESS AND HEALTHINESS ITEMS FOR EACH FOOD IN STUDY 1B AFTER VARIMAX ROTATION

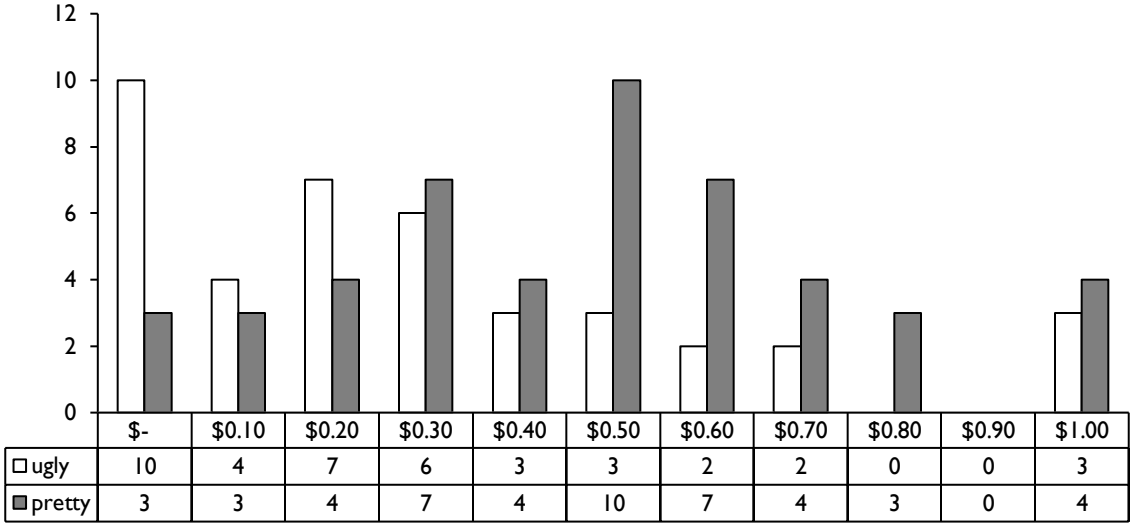
Item	Component		
Almond butter & banana toast	1	2	3
nutritious	.900	.252	.089
good for me	.883	.238	.191
healthy	.876	.283	.197
pure	.269	.856	.134
natural	.314	.848	.140
unprocessed	.181	.830	.229
low calorie	.164	.170	.902
low fat	.162	.191	.897
Spaghetti marinara	1	2	
nutritious	.888	.183	
good for me	.882	.260	
healthy	.876	.277	
low in calories	.576	.345	
low in fat	.501	.314	
unprocessed	.217	.871	
pure	.306	.855	
natural	.333	.855	
Cupcake	1	2	
healthy	.824	.320	
low in calories	.823	.223	
nutritious	.814	.304	
good for me	.797	.295	
low in fat	.787	.244	
unprocessed	.192	.849	
natural	.365	.840	
pure	.339	.840	

F: STUDY 2 SUPPLEMENTARY ANALYSES

Results

WTP. The distribution of WTP is displayed in Figure S3. Analyzing the data in terms of propensity to bid (vs. not bid) instead of willingness-to-pay, reveals that people were also significantly more likely to bid something (rather than nothing) on the pretty pepper (46 out of 49, 93.9%) than on the ugly pepper (30 out of 40, 75.0%), $\chi_2(1) = 6.29$, $p = .012$.

FIGURE S3
WTP DISTRIBUTION IN STUDY 2B BY CONDITION



	$-	$0.10	$0.20	$0.30	$0.40	$0.50	$0.60	$0.70	$0.80	$0.90	$1.00
ugly	10	4	7	6	3	3	2	2	0	0	3
pretty	3	3	4	7	4	10	7	4	3	0	4

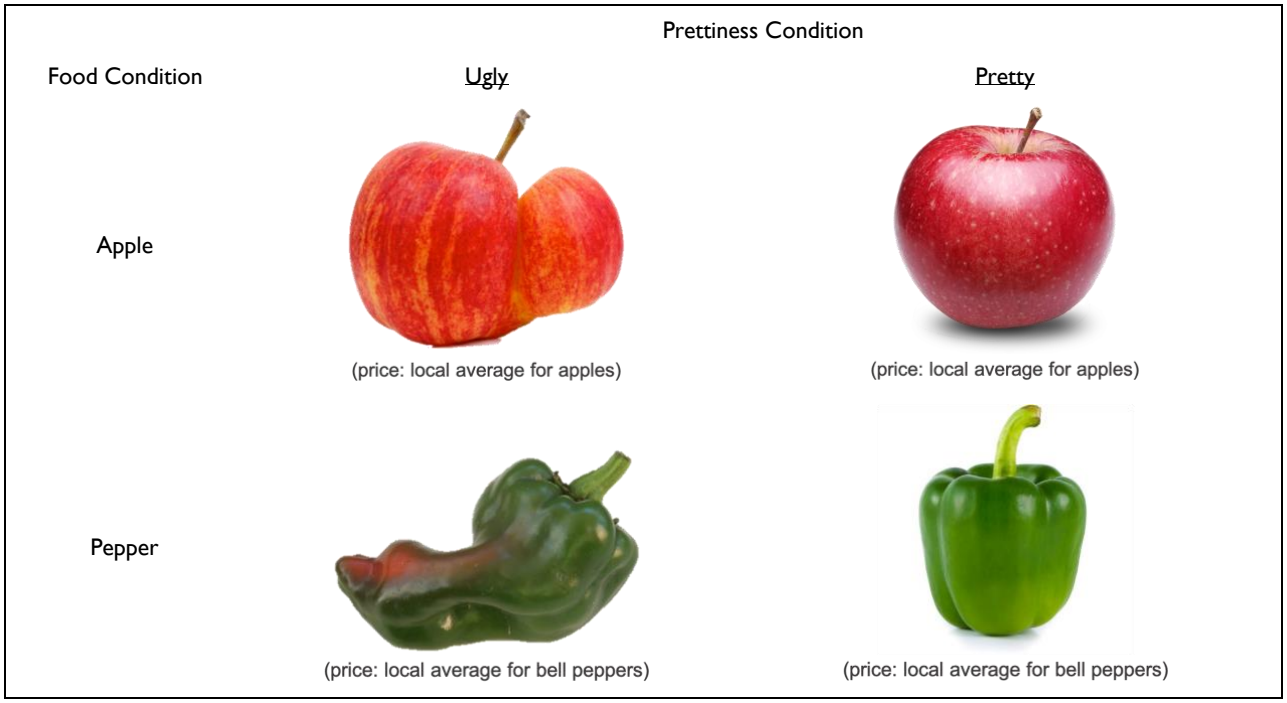
G: STUDY 2 CONCEPTUAL REPLICATION: PRETTINESS BOOSTS WILLINGNESS-TO-PAY VIA NATURALNESS AND HEALTHINESS

This experiment replicated the patterns we found in study 2 using real live bell peppers with pictures of two different kinds of produce and hypothetical willingness-to-pay (WTP). It also shows mediation via perceived naturalness.

Method

Two hundred Amazon Mechanical Turk panelists in the United States (51.0% women; M_{age} = 35.01, $range_{age}$ = 19–68) were randomly assigned to evaluate a photograph of an apple or a bell pepper that was either pretty or ugly (see Figure S4), creating a 2(prettiness: pretty vs. not pretty)×2(produce: apple vs. pepper) design. We used two kinds of produce for robustness but did not predict differences by produce type.

FIGURE S4
STIMULI USED IN THE CONCEPTUAL REPLICATION OF STUDY 2



As theories of classical aesthetics predict, in a pretest (see Web Appendix B), people had viewed the apple as significantly prettier when it was symmetrical and evenly balanced (M = 6.08, SD = 1.07) than when it was asymmetrical and less balanced (M = 3.59, SD = 1.89; $p < .001$), and the pepper as significantly prettier when it was symmetrical and had a clear, uniform color (M = 5.86, SD = 1.25) than when it was asymmetrical and had blurred multiple colors (M = 2.69, SD = 1.65; $p < .001$).

Participants received similar instructions as in studies 1A–B and the replication. Below each image, participants read that the price was the "local average." Then they completed, in randomized order, the same "perceived healthiness" scale ($\alpha = .91$), "perceived naturalness" scale ($\alpha = .88$), and "perceived tastiness" scale ($\alpha = .96$). (Note that principal component analysis showed that prettiness and healthiness are statistically distinct constructs, suggesting discriminant validity; see below.)

Willingness-to-Pay. Then we measured willingness-to-pay (WTP) for the piece of produce. Participants chose a reservation price between $0 and $2 (10¢ increments).

Finally, they reported demographics, which included education level and household income. Neither interacted with prettiness on perceived healthiness or naturalness, so these variables are not discussed further (although details and results are available upon request).

Results and Discussion

Basic effect: Perceived healthiness. There was only a main effect of prettiness, such that participants rated the produce as significantly healthier when it was pretty (M = 6.30, SD = .93) than when it was ugly (M = 5.46, SD = 1.28; $F(1, 196) = 27.81$, $p < .001$, d = .75). There was no effect of produce ($F(1, 196) = .62$, $p = .430$) and no prettiness×produce interaction ($F(1, 196) = .81$, $p = .369$).

Proposed mediator: Perceived naturalness. There was only a main effect of prettiness, such that participants rated the produce as significantly more natural when it was pretty (M = 6.22, SD = .97) than when it was ugly (M = 5.35, SD = 1.74; F(1, 196) = 19.56, $p < .001$, d = .62). There was no effect of produce (F(1, 196) = 1.17, $p = .280$) and no prettiness×produce interaction (F(1, 196) = .00, $p = .99$).

Perceived tastiness. There was a main effect of prettiness, such that participants rated the produce as significantly tastier when it was pretty (M = 5.53, SD = 1.40) than when it was ugly (M = 4.23, SD = 1.89; F(1, 196) = 31.32, $p < .001$, d = .78). There was also an incidental main effect of produce, such that participants rated the apple as significantly tastier (M = 5.23, SD = 1.53) than the pepper (M = 4.54, SD = 1.95; F(1, 196) = 8.64, $p = .004$, d = .39), but no prettiness×produce interaction (F(1, 196) = .62, $p = .434$).

WTP. There was a main effect of prettiness, such that participants were willing to pay significantly more for the produce when it was pretty (M = $.86, SD = $.47) than when it was ugly (M = $.53, SD = $.39; F(1, 196) = 31.01, $p < .001$, d = .76). There was also an incidental main effect of produce, such that participants were willing to pay significantly more for the pepper (M = $.76, SD = $.51) than the apple (M = $.63, SD = $.39; F(1, 196) = 4.58, $p = .034$, d = .29), and an incidental main effect prettiness×produce interaction (F(1, 196) = 5.54, $p = .020$), such that the difference in WTP was more pronounced for the pepper ($M_{diff} = .47$, SE = .08; $p < .001$) than the apple ($M_{diff} = .19$, SE = .08; $p = .024$).

Mediation. We tested several mediation models. Parallel mediation (model 4; 10,000 samples) with perceived healthiness and tastiness simultaneously indicated that both perceived healthiness (B = .024, SE = .013; 95% CI [.004, .054]) and perceived tastiness (B = .042, SE = .014; 95% CI [.019, .075]) mediated the effect of prettiness on WTP, each contributing above

and beyond the other.

More pertinent to our hypotheses, serial mediation (model 6; 10,000 samples) with perceived naturalness and perceived healthiness showed that a naturalness–healthiness sequence mediated the effect of prettiness on WTP (B = .018, SE = .007; 95% CI [.007, .037). Conversely, the opposite healthiness–naturalness sequence did not (B = .007, SE = .007; 95% CI [–.007, .022). Additionally, replicating and expanding the results from the field study in the main text, a naturalness–healthiness–tastiness sequence where healthiness gives rise to tastiness mediated (B = .009, SE = .004; 95% CI [.003, .017]), but a naturalness–tastiness–healthiness sequence did not mediate (B = .002, SE = .001; 95% CI [–.003, .005]) and, likewise, a tastiness–naturalness–healthiness sequence also did not mediate (B = .003, SE = .002; 95% CI [–.001, .009]).

Principal component analysis. Principal component analysis indicated that a two-factor solution best represented the eight items: two factors had eigenvalues above 1 (5.07 and 1.16), and the second factor explained a considerable amount of variance (14.54%) beyond what the first factor explained (63.32%). The factor loadings are displayed in Table S8. The "perceived naturalness" items loaded together on one factor; the "perceived healthiness" items on the other.

TABLE S8
FACTOR LOADINGS OF HEALTHINESS AND PRETTINESS ITEMS IN THE CONCEPTUAL REPLICATION OF STUDY 2 AFTER VARIMAX ROTATION

Item	Component	
	1	2
pure	.886	.246
natural	.844	.317
unprocessed	.805	.083
low in calories	.081	.881
low in fat	.182	.846
nutritious	.520	.705
good for me	.597	.670
healthy	.625	.648

H: SUPPLEMENTAL STUDY: COGNITIVE NATURE OF THE PRETTY=NATURAL=HEALTHY ASSOCIATIONS

Going beyond visual stimuli, we conducted a supplemental study to explore how the pretty=natural and natural=healthy associations psychologically operate on the explicit and implicit levels. The cognitive nature of the associations has key implications for potential moderators. For instance, if the effect is based on strong implicit associations, interventions targeting the explicit level (e.g., verbal information) would likely require very salient information and deliberation to override the implicit associations. We explore each association (i.e., pretty=natural and natural=healthy) with a different sub-sample, but use the same general procedure for both.

Method (Pretty=Natural Association)

One hundred and two Amazon Turk panel members (45.1% women; M_{age} = 39.75, $range_{age}$ = 22–75) first performed an Implicit Association Test (IAT; Greenwald, McGhee, and Schwartz 1998) and then rated their explicit beliefs. The IAT measures reaction time via key strokes, so participation was restricted to desktop computers.

Implicit Association Test procedure. The IAT is the most commonly used instrument to measure implicit associations. We administered its survey-based version for Qualtrics (Carpenter et al. 2019). In an IAT, participants must categorize words as quickly and accurately as they can. If people implicitly associate two concepts (e.g., pretty=natural), they will categorize a target stimulus (e.g., "pure") as part of a pair of categories more quickly when these categories are compatible (e.g., does "pure" belong with natural/pretty or artificial/ugly?) compared to when they are incompatible (e.g., does "pure" belong with natural/ugly or artificial/pretty?). Per

standard procedures (Greenwald et al. 1998), participants performed seven blocks: five practice blocks (categorization into individual categories, e.g., is [target] pretty or ugly; is [target] natural or artificial?) and two counterbalanced test blocks (categorization into category pairs)—one compatible (e.g., is [target] natural/pretty or artificial/ugly?) and the other incompatible (e.g., is [target] natural/ugly or artificial/pretty?). The difference in average response time between the compatible and the incompatible block quantifies the strength of the association.

IAT for the pretty=natural association. For the pretty=natural association, we used the categories pretty, ugly, natural, and artificial; target stimuli were words such as handsome, hideous, pure, and refined (see Table S9 for a list of target words).

TABLE S9
CATEGORIES AND TARGET WORDS IN IMPLICIT ASSOCIATION TESTS

	Pretty=Natural					
			Natural=Healthy			
Concept	Prettiness		Naturalness		Healthiness	
Categories	pretty	ugly	natural	artificial	healthy	unhealthy
Target words	attractive	bad-looking	organic	altered	fit	fatty
	beautiful	deformed	pristine	refined	good for me	harmful
	good-looking	hideous	pure	synthetic	lean	ill
	gorgeous	unattractive	raw	tainted	nourishing	poisonous
	handsome	unsightly	unprocessed	transformed	nutritious	sickening

Explicit pretty=natural belief. Participants then rated their agreement with three statements: "Beautiful foods usually look natural;" "Foods that are pretty tend to be highly processed;" and "The better food looks, the more pure it is." They provided ratings on 7-point scales anchored at 1 = completely disagree and 7 = completely agree. The second item was reverse-scored and all three items were combined into a "pretty=natural belief" scale (α = .69).

Results (Pretty=Natural Association)

Implicit Association Test. Data were analyzed following prevailing validated guidelines (Greenwald, Nosek, and Banaji 2003). First, trials with excessive reaction time (>10,000ms;

.16% of all 20,400 trials) and participants with a high proportion of too short reaction times (>10% <300ms; 7.84% of all 102 participants) were dropped (although results remain significant when these trials and participants are retained). Then, a standardized difference score (D-score) representing the difference in response time between the incompatible and the compatible block was computed for each participant. Higher D-scores indicate people responded faster to the compatible block than the incompatible block; D-scores of 0 indicate no difference in speed.

A one-sample t-test showed that people responded significantly faster to the compatible block (i.e., is [target] natural/pretty or artificial/ugly?) compared to the incompatible block (i.e., is [target] natural/ugly or artificial/pretty?; D = .75, SD = .42; t(93) = 17.33, $p < .001$, d = 1.79).

Explicit pretty=natural belief. Explicit pretty=natural belief scores (M = 4.08, SD = 1.24) did not differ from the neutral scale midpoint (4) (t(101) = .61, $p = .54$, d = .06).

D-scores from the IAT were significantly greater than zero across all four quartiles of strength of the explicit pretty=natural belief (ts > 6.98, *p*s < .001). That is, people possess an implicit association between prettiness and naturalness irrespective of their explicit intuition.

Method (Natural=Healthy Association)

A different sub-sample of 100 Amazon Turk panel members (49.00% women; M_{age} = 38.35, $range_{age}$ = 18–72) completed the same tasks as described above with material pertaining to the natural=healthy association.

IAT for the natural=healthy association. For the natural=healthy association, we used the categories natural, artificial, healthy, and unhealthy; targets were words such as pure, refined, nutritious, and ill (see Table S9 for a list of target words).

Explicit natural=healthy belief. Participants rated to what extent they agreed with three statements: “The healthiest foods are the most natural foods;” “Processed foods tend to be less

nutritious;" and "All-natural foods are usually better for our body." All three items were combined into a "natural=healthy belief" scale (α = .84).

Results (Natural=Healthy Association)

Implicit Association Test. Data were analyzed as described above. Trials with excessive reaction time (>10,000ms; 0.13% of all 20,000 trials) and participants with a high proportion of too short reaction times (<300ms in >10% of trials; 8.0% of all 100 participants) were dropped (though results remain significant when these trials and participants are retained). A one-sample t-test showed that people responded significantly faster to the compatible block (i.e., is [target] natural/healthy or artificial/unhealthy?) compared to the incompatible block (i.e., is [target] natural/unhealthy or artificial/healthy?; D = .87, SD = .40; t(91) = 20.77, $p < .001$, d = 2.17).

Explicit natural=healthy belief. Explicit natural=healthy belief scores (M = 5.83, SD = 1.14) were heavily left-skewed and leptokurtic (Md = 6.00; M = 5.83, SD = 1.44; Shapiro-Wilk W (92) = .857, $p < .001$). This distribution violates the normality assumption—a key requirement for parametric tests. Thus, instead of a parametric one sample t-test we applied its non-parametric equivalent, a one-sample Wilcoxon ranked sign test (median test), to test belief scores against the scale midpoint. Natural=healthy belief scores were significantly above the scale midpoint (4) (W+ = 4035, z = 7.71, $p < .001$). Results remain unchanged when using a parametric test (t(91) = 15.30, $p < .001$, d = 1.61).

D-Scores from the IAT were significantly greater than zero across all four quartiles of strength of the explicit natural=healthy belief (ts > 8.43, ps < .001). That is, people possess an implicit association between naturalness and healthiness irrespective of their explicit intuition.

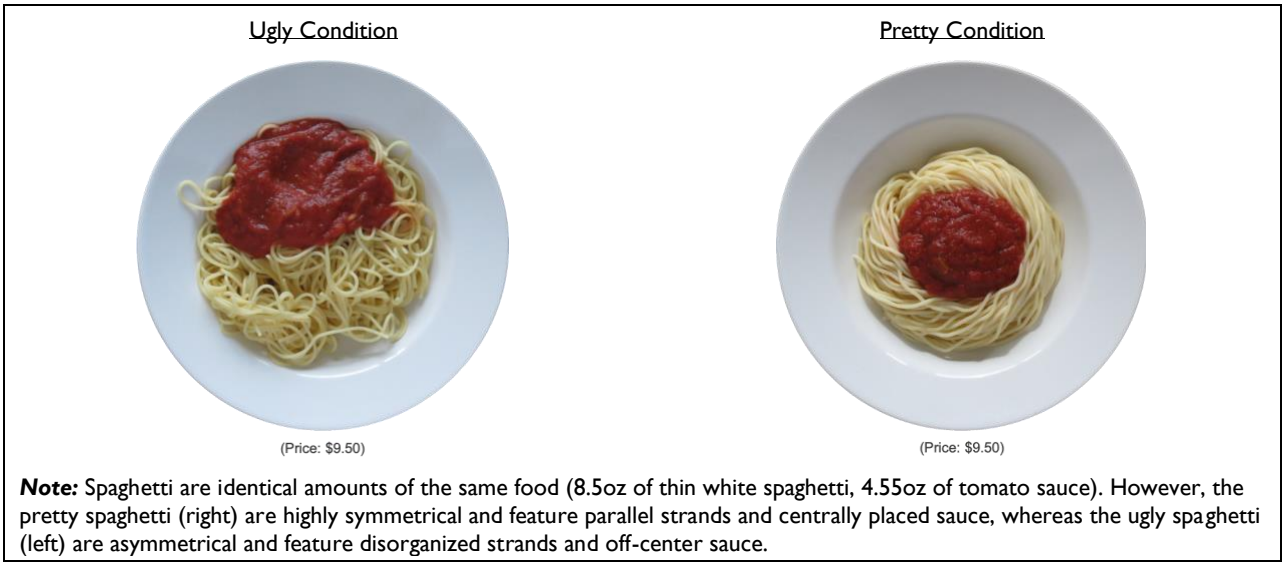
I: SUPPLEMENTAL MODERATION STUDY: STRONGER MEDIATING EFFECT AMONG THOSE WITH A STRONGER NATURAL=HEALTHY BELIEF

This supplemental study tests the predicted moderation by belief in natural=healthy. If perceived naturalness drives the effect, theoretically, the mediating effect of naturalness should be stronger for people with a stronger belief in the natural=healthy relationship. Practically, such a moderation would imply that those consumers who are likely the most interested in choosing natural foods for health reasons are, ironically, also more susceptible to the pretty=healthy effect.

Method

Four hundred and three Prolific Academic panel members in the United States (51.9% women; M_{age} = 33.44, $range_{age}$ = 18–75) were randomly assigned to evaluate a plate of spaghetti marinara that looked either pretty or ugly (see Figure S5), before we measured their belief in the natural=healthy relationship.

FIGURE S5
STIMULI USED IN ADDITIONAL MODERATION STUDY



Note: Spaghetti are identical amounts of the same food (8.5oz of thin white spaghetti, 4.55oz of tomato sauce). However, the pretty spaghetti (right) are highly symmetrical and feature parallel strands and centrally placed sauce, whereas the ugly spaghetti (left) are asymmetrical and feature disorganized strands and off-center sauce.

As theories of classical aesthetics predict, in a pretest (see Web Appendix B), people had found the spaghetti marinara (8.5oz of spaghetti, 4.55oz of tomato sauce) significantly prettier when they featured a high degree of symmetry and order (M = 4.50, SD = 1.56) than when they did not (M = 2.68, SD = 1.67; $p < .001$). The same price of $9.50 was shown under each image.

Participants received the same instructions and completed the same "perceived healthiness" ($\alpha = .84$), "perceived naturalness" ($\alpha = .90$), and "perceived tastiness" ($\alpha = .94$) scales as in prior studies. Additionally, we measured perceived freshness and amount of food.

Perceived freshness. They rated to what extent the food looked fresh and stale. The latter was reverse-coded and both combined into a "perceived freshness" composite ($r = .47$, $p < .001$).

Perceived amount. They also rated to what extent the food looked filling, satiating, large, and substantial. All three items were combined into a "perceived amount" scale ($\alpha = .90$).

They also rated the "perceived price" measure (r = .83, $p < .001$).

Proposed moderator: Belief in natural=healthy. Lastly, we measured people's belief in the natural=healthy relationship using the "natural=healthy belief" scale ($\alpha = .78$) as in the supplemental study (Web Appendix H).

Recall that, conceptually, we only predict belief in natural=healthy to moderate the effect of the mediator on the dependent variable (path b). There is no reason to expect belief in natural=healthy to have any main effect on perceived healthiness or to moderate the direct effect of prettiness on perceived healthiness. Additionally, from an empirical standpoint, an interaction between prettiness and belief in natural=healthy on perceived healthiness is very unlikely to emerge. The supplemental study (Web Appendix H) revealed strong implicit associations in beyond the explicit beliefs for both mediating paths (a and b), so the natural=healthy belief is just one of four components in the underlying process. Finally, participants reported demographics.

Results and Discussion

Basic effect: Perceived healthiness. Participants rated the spaghetti as significantly healthier when they were pretty (M = 3.71, SD = 1.14) than when they were ugly (M = 3.45, SD = 1.18; F(1, 401) = 4.92, p = .027, d = .22).

Proposed mediator: Perceived naturalness. Participants rated the spaghetti as significantly more natural when they were pretty (M = 3.45, SD = 1.45) than when they were ugly (M = 3.04, SD = 1.42; F(1, 401) = 8.17, p = .004, d = .29).

Perceived tastiness, freshness, and amount. Participants rated the spaghetti as significantly tastier when they were pretty (M = 3.98, SD = 1.59) than when they were ugly (M = 3.56, SD = 1.65; F(1, 401) = 6.74, p = .010, d = .26). They also rated the spaghetti as fresher when they were pretty (M = 4.81, SD = 1.38) than when they were ugly (M = 4.02, SD = 1.53; F(1, 401) = 29.51, p < .001, d = .54) and the amount as larger when they were pretty (M = 4.17, SD = 1.37) than when they were ugly (M = 3.89, SD = 1.42; F(1, 401) = 4.26, p = .040, d = .20).

Equality check for perceived price. As intended, participants' price perceptions were virtually equal, regardless of whether the spaghetti were very pretty (M = 4.49, SD = 1.97) or ugly (M = 4.33, SD = 2.14; F(1, 401) = .62, p = .430, d = .08).

Proposed moderator: Natural=healthy belief. Belief scores were again non-normally distributed (Shapiro-Wilk W (403) = .941, p < .001), so we applied a Mann-Whitney U test. People held equally strong natural=healthy beliefs in the pretty (Md = 5.67; M = 5.40, SD = 1.22) and the ugly condition (Md = 5.67; M = 5.48, SD = 1.12; U = 19899, z = –.346, p = .730). Results remain unchanged when using a parametric test (F(1, 401) = .40, p = .53, d = .07).

Mediation. Moderated mediation (model 14; 10,000 samples) with perceived naturalness as the mediator and belief in natural=healthy moderating the path from mediator to the dependent

variable (path b) returned a significant interaction between perceived naturalness and belief natural=healthy on perceived healthiness (B = .058, SE = .027, p = .033) and a significant index of moderated mediation (B = .024, SE = .016; 95% CI [.002, .067]). Specifically, perceived naturalness had a stronger mediating effect of prettiness on perceived healthiness for people with a strong belief in natural=healthy (B_{+1SD} = .220, SE = .077; 95% CI [.076, .373]) compared to people with a weak belief (B_{-1SD} = .164, SE = .061; 95% CI [.056, .297]).

We also tested parallel mediation (model 4; 10,000 samples) with naturalness and freshness simultaneously. Central to testing our theory about the role of perceived naturalness, the mediation pattern for naturalness remained unchanged. In a simple parallel mediation, naturalness mediated the effect of prettiness on perceived healthiness (B = .188, SE = .067; 95% CI [.063, .324]), whereas freshness did not (B = .039, SE = .033; 95% CI [–.019, .109]).

Note that there is no reason to predict the natural=healthy belief to moderate the path from perceived freshness to perceived healthiness, but we also tested moderated mediation (model 14; 10,000 samples) with naturalness and freshness as joint mediators and belief in natural=healthy as a moderator. Again, naturalness mediated (B_{+1SD} = .211, SE = .075; 95% CI [.060, .359]; B_{-1SD} = .156, SE = .060; 95% CI [.043, .284]), whereas freshness did not (B_{+1SD} = .045, SE = .041; 95% CI [–.029, .134]; B_{-1SD} = .038, SE = .042; 95% CI [–.039, .129]). (As suspected, there was no interaction between perceived freshness and belief in natural=healthy on perceived healthiness, B = .004, SE = .027, p = .882.)

As the conceptual model predicts, the strength of the natural=healthy association moderates the mediating effect of naturalness. However, as expected in light of the strong implicit natural=healthy association, a low explicit natural=healthy belief is not sufficient to eliminate the pretty=healthy effect.

J: STUDY 4A SUPPLEMENTARY METHODS AND ANALYSES

Method

Perceived sophistication. Participants rated to what extent they thought to food was sophisticated, fancy, and ordinary$_R$. The latter was reverse-scored and all three items were combined into a "perceived sophistication" scale (α = .76).

Perceived care. They also rated to what extent they thought the food was carefully made, thoughtfully sourced, and handmade, which were combined into a "perceived care" scale (α = .63).

Results

Perceived sophistication. Condition had a significant effect ($F(2, 598) = 76.83$, $p < .001$). Compared to the ugly condition (M = 2.70, SD = 1.25), people rated the toast significantly more sophisticated in the pretty–classical condition (M = 3.40, SD = 1.39; $t(598) = –5.30$, $p < .001$, d = .53), and also more sophisticated in the pretty–expressive condition (M = 4.33, SD = 1.33; $t(598) = –12.36$, $p < .001$, d = 1.26). Importantly, they rated the toast significantly *less* sophisticated in the pretty–classical than in the pretty–expressive condition ($t(598) = 7.03$, $p < .001$, d = .68).

Perceived care. Condition had a significant effect ($F(2, 598) = 42.06$, $p < .001$). Compared to the ugly condition (M = 4.72, SD = 1.07), people rated the toast as significantly more carefully made in the pretty–classical condition (M = 5.53, SD = 1.06; $t(598) = –7.58$, $p < .001$, d = .76) and also more carefully made in the pretty–expressive condition (M = 5.60, SD = 1.07; $t(598) = –8.25$, $p < .001$, d = .82). Importantly, they rated the toast as equally carefully made in the pretty–classical and pretty–expressive conditions ($t(598) = .66$, $p = .508$, d = .07).

Additional Mediation. We also conducted parallel mediation with perceived naturalness,

care, and sophistication simultaneously (multi-categorical independent variable with indicator coding, ugly serving as the reference category; model 4; 10,000 samples).

First and foremost, all mediation patterns held for naturalness. For the comparison between the ugly and the pretty–classical aesthetics conditions, naturalness mediated the effect of prettiness on perceived healthiness (B = .237, SE = .057; 95% CI [.129 to .351]). For the comparison between the ugly and the pretty–expressive conditions, naturalness did not mediate (B = –.045, SE = .058; 95% CI [–.160 to .066])—as expected, given that these two groups differed on neither naturalness nor healthiness. An additional mediation with pretty–expressive as the reference category showed that for the comparison between the pretty–expressive and the pretty–classical conditions, naturalness also mediated the effect of prettiness on perceived healthiness (B = .282, SE = .060; 95% CI [.170 to .405]).

In addition to naturalness, care had some explanatory value. For the comparison between the ugly and the pretty–classical aesthetics conditions, care mediated part of the effect of prettiness on perceived healthiness (B = .193, SE = .047; 95% CI [.107 to .291]), albeit less so than naturalness (see above). For the comparison between the ugly and the pretty–expressive conditions, care also mediated (B = .210, SE = .050; 95% CI [.119 to .316])—unexpected, given that these groups did *not* differ on healthiness. An additional mediation with pretty–expressive as the reference category showed that for the comparison between the pretty–expressive aesthetics and the pretty–classical aesthetics conditions, care did *not* mediate (B = –.017, SE = .026; 95% CI [–.070 to .033])—as expected, given that these two groups did not differ on care.

Sophistication did not mediate for any comparison (ugly vs. pretty–classical 95% CI [–.007 to .079]; ugly vs. pretty–expressive 95% CI [–.017 to .177]; pretty–expressive vs. pretty–classical 95% CI [–.106 to .010]).

K: STUDY 4B SUPPLEMENTARY METHODS AND ANALYSES

Method

Perceived sophistication. Participants rated how "large" and "filling" the food seemed, which were combined into a "perceived amount" composite (r = .58, p < .001).

Results

Equality checks for perceived tastiness and price. As intended, condition had no significant effect on perceived tastiness (F(2, 298) = 1.23, p = .294) or perceived price (F(2, 298) = .89, p = .410).

Perceived amount. Condition had a significant effect (F(2, 298) = 9.85, p < .001). Compared to the ugly condition (M = 3.49, SD = 1.28), people rated the toast as significantly larger in the pretty condition (M = 4.23, SD = 1.42; t(298) = –3.75, p < .001, d = .55), and also significantly larger in the pretty+disclaimer condition (M = 4.27, SD = 1.49; t(298) = –3.93, p < .001, d = .56). They rated the toasts as equally large in the pretty and the pretty+disclaimer conditions (t(298) = .189, p = .851, d = .10).