Journal of Experimental Psychology: General

Fleeting Impressions of Economic Value via Summary Statistical Representations

Allison Yamanashi Leib, Kelly Chang, Ye Xia, Andy Peng, and David Whitney

Online First Publication, April 27, 2020. http://dx.doi.org/10.1037/xge0000745

American Psychological Association
Journal of Experimental Psychology: General

ISSN: 0096-3445
http://dx.doi.org/10.1037/xge0000745

Fleeting Impressions of Economic Value via Summary Statistical Representations

Allison Yamanashi Leib, Kelly Chang, Ye Xia, Andy Peng, and David Whitney
University of California, Berkeley

Visual processing is limited: we cannot exhaustively analyze every object in a scene in a brief glance. However, ensemble perception affords the visual system a rapid shortcut to efficiently evaluate multiple objects. Ensemble processing has been widely tested across basic features. However, ensemble perception could be especially important and valuable for processes that are normally thought to require cognitive deliberative effort. One typical high-level cognitive process that humans engage in frequently is evaluating the value of objects. Here, we presented brief displays of consumer products to human observers and measured their visual sensitivity to the average value of the sets. We found that participants were sensitive to the average value of sets of products even when they did not have explicit memory for every item in the display. Our results show that value judgments can be based on ensemble information. Although value is thought to be an inferential concept, ensemble processing affords the brain a heuristic to efficiently assign value to entire sets of objects.

Keywords: ensemble perception, price, consumer behavior, vision

Supplemental materials: http://dx.doi.org/10.1037/xge0000745.supp

Many visual impressions are formed at the level of perceptual sets. For example, an individual walking down the street can easily evaluate the ensemble emotions of nearby pedestrians (Elias, Dyer, & Sweeny, 2017; Haberman & Whitney, 2007, 2009), the ensemble speed of passing cars (Watamaniuk & Duchon, 1992), or the ensemble hue of surrounding trees (Demeyere, Rzeskiewicz, Humphreys, & Humphreys, 2008; Webster, Kay, & Webster, 2014). These ensemble percepts, or statistical descriptions of groups, remain robust even when detailed information about specific members in the set are unavailable to the visual system (Alvarez, 2011; Hochstein, Pavlovskaya, Bonneh, & Soroker, 2015; Whitney & Yamanashi Leib, 2018). As such, ensemble percepts are functionally useful heuristics affording critical information about our visual environment that cannot be otherwise accessed.

Previous literature demonstrates that observers can successfully evaluate ensemble information from concrete visual features such as: orientation (Parkes, Lund, Angelucci, Solomon, & Morgan, 2001), hue (Demeyere et al., 2008; Webster et al., 2014), motion direction (Watamaniuk & McKee, 1998,Watamaniuk, Sekuler, & Williams, 1989), speed (Watamaniuk & Duchon, 1992), size (Ariely, 2001; Chong, Joo, Emmanouil, & Treisman 2008; Chong & Treisman, 2003, 2005; Marchant & de Fockert, 2009; Marchant, Simons, & de Fockert, 2013), facial expression (Elias, et al., 2017; Haberman & Whitney, 2007, 2009), biological motion (Sweeny, Haroz, & Whitney, 2013), gaze (Florey, Clifford, Dakin, & Mareschal, 2016; Florey, Dakin, & Mareschal, 2017; Sweeny & Whitney, 2014), and other high level visual information like attractiveness (Post, Haberman, Iwaki, & Whitney, 2012; van Osch, Blanken, Meijs, & van Wolferen, 2015; Walker & Vul, 2014) and animacy (Yamanashi Leib, Kosovicheva, & Whitney, 2016). Ensemble representations are obviously valuable in these cases because they provide a short-cut to rapidly evaluate similar or redundant objects in the surrounding environment (Alvarez, 2011), especially when there is not sufficient time or visual capacity to exhaustively analyze each person or object in a scene (Whitney & Yamanashi Leib, 2018). However, it remains an open question whether ensemble representations are formed for other complex but commonly encountered visual stimuli that require abstraction, such as the value of consumer products. The perceived value of products in store windows is not directly conveyed by explicit, redundant visual features (unlike the hue of leaves on a tree or the facial expressions in a crowd, where there are similar or redundant visual features; Zeithaml, 1988). Most importantly, the ensemble value of products does not vary in conjunction with simple visual features across the stimuli set; rather, visual features associated with product quality vary immensely between product categories (Bonner & Nelson, 1985; Laird, 1932; Olshavsky, 1985). Indeed, a visual feature that is correlated with quality in one product category can be anticorrelated with quality in another product

Allison Yamanashi Leib, Kelly Chang, Ye Xia, Andy Peng, and David Whitney, Department of Psychology, University of California, Berkeley.

Kelly Chang is now at the Department of Psychology, University of Washington. Ye Xia is now at Google, Mountain View, California. Andy Peng is now at Computer Science Department, University of California, Los Angeles.

We thank Mateo Lopez for data analysis.

The authors declare no competing financial interest.

Data is available upon request.

Correspondence concerning this article should be addressed to Allison Yamanashi Leib, who is now at Fitbit, 199 Fremont Street, San Francisco, CA 94105. E-mail: ayleib@berkeley.edu

category (Zeithaml, 1988). For example, large size is often an indicator of higher price for common foods (e.g., a larger container of cereal often costs more than a smaller container). However, simple heuristics fall apart quickly since generic products can be larger and cheaper, and size is anticorrelated with quality for electronic devices (e.g., smaller, sleeker designs are often associated with increased price). Therefore when judging perceived price across stimuli from different product categories, individuals must *abstract* the ensemble value from a collection of disparate physical features. They must go one step further, beyond a basic feature analysis, to access a gist impression of ensemble product value. Despite a great deal of research on the perception of individual object price (e.g., Bishop, 1984; Doyle, 1984; Jacoby & Olson, 1985; Sawyer & Dickson, 1984; Schechter, 1984), it is less clear if and how product sets, as a whole, are evaluated and how accurate or fast this process is. The majority of consumer product research focuses on later stages of product purchasing, rather than the early stages when consumers may be first drawn to a product display in a split-second glance. Rapid viewing of product sets is ubiquitous in consumer environments (e.g., scrolling past ads on the computer, riding past ads on the subway, walking past fruit stands at a farmer's market, etc.). Because sets of products are the default form in which products are encountered, at least initially, it is particularly relevant to investigate observers' first-glance impressions of product sets. Ultimately, purchasing may still require longer cognitive deliberation. However, if snap-judgment ensemble valuations are made by observers, these could be an equally critical component of consumer behavior. For example, ensemble valuations may be a significant factor in consumers' decisions to stop scrolling or pause walking and initiate shopping. In the following experiments, we tested whether observers can extract average value information from product sets presented during a brief glance—in a second or less.

If observers are able to evaluate ensemble value of a set of objects efficiently, this could be a useful function for strategic shopping, guiding search for particular objects, and for forming preferences. In addition to having practical implications for consumer purchasing, investigations such as these can inform us about the nature of ensemble perception itself. Visual sensitivity to ensemble value would confirm that ensemble perception is not only an expedient visual mechanism for identifying redundant concrete features but may also provide a fast shortcut to abstract perceptual impressions.

Experiment 1a: Ensemble Value of Product Sets

Method

Participants. Experiment 1a included 90 observers recruited from Amazon Mechanical Turk (47 females, 43 males). The sample size was based on prior published work investigating ensemble perception using an entirely online population (Goodale, Alt, Lick, & Johnson, 2018). We asked participants to proceed with the experiments only if they had no neurological history and normal or corrected-to-normal vision. All participants were consented in accordance with the Institutional Review Board at UC Berkeley. All methods and experimental procedures were approved by the Institutional Review Board at UC Berkeley.

Stimuli. The stimuli were 100 pictures of consumer products that were downloaded from the website of a major retailer. The stimulus array included consumer products typically found at major retail chains such as: food, clothing, office supplies, travel items, electronics, and so forth Examples of the products are shown in Figure 1. The least expensive product in the stimulus array was listed by the retailer at \$0.72, while the most expensive product was listed by the retailer at \$99.99. Observers were not given any information about the individual prices or the range of product prices. We did not remove brand names from the products. As such, participants who were highly familiar with the retailer may have guessed the retailer by associating specific brands with the store. However, the specific name of the retailer was not provided to participants. There were 100 products in the stimulus array. In each trial, a set of six products was presented on a white background in a 2 × 3 grid that measured 453 × 654 pixels. The experiments were conducted on participants' home computers; as such, the visual angle and the luminance of the stimuli cannot be reported. Stimuli was displayed using the Qualtrics (Qualtrics, Copyright 2018) experimental platforms.

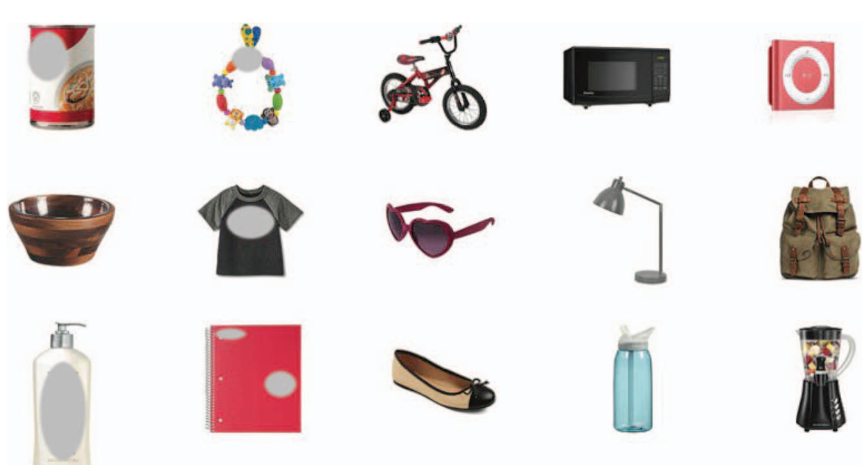

Figure 1. Examples of stimuli. Fifteen (out of 100) stimuli in the products used in Experiments 1, 2, and 3. See the online article for the color version of this figure.

Procedure. Participants were asked to place their personal computer monitor in a centered position in front of them, and were asked to maintain a clear, unobstructed view of the screen. They were also asked to sit an arm's length away from the computer screen. They were instructed that displays would be brief, and they were encouraged to maintain vigilant attention throughout the experiment. First, participants viewed a set of 6 products for one second. The 6 products were drawn pseudorandomly, with the restriction that no set contained a duplicate product. We also included a constraint that the average value of the set had to vary (as a true random draw often results in a repeated mean value). The expected average value of each set of stimuli was determined by calculating the arithmetic mean of all 6 products contained in the set, derived from the retailer's listed price. From now on, we will refer to this as the "predicted value" for each set of products. Some product sets contained less than 6 products (these conditions are explained in detail below). In product sets containing less than 6 products, the products were randomly distributed across the six possible locations on the grid. All product displays contained a centrally placed fixation cross. Participants were asked to keep their eyes fixated on the cross while the product set was on the screen. After the products disappeared, participants entered their estimate of the average price of the product set in a blank text box that appeared centrally on the screen. Participants were given an unlimited time to report their price estimates. They were not allowed to proceed to the next trial until they entered a response. See Figure 2A for an illustration of the trial sequence. Observers participated in 100 trials total. On 25 of these trials, there were 6 products displayed in the set. In the remaining 75 trials, there were displays of 4, 2, or 1 product, in equal proportion. These subsets of the entire 6 product set allowed us to later perform an analysis we will call the subset integration measure, which estimates how many stimuli observers incorporated into their ensemble percepts (Chong et al., 2008; Piazza, Sweeny, Wessel, Silver, & Whitney, 2013; Sweeny, Haroz, & Whitney, 2013; Sweeny & Whitney, 2014; Sweeny, Wurnitsch, Gopnik, & Whitney, 2015; Wolfe, Kosovicheva, Leib, Wood, & Whitney, 2015; Yamanashi Leib et al., 2014; Yamanashi Leib, Kosovicheva, & Whitney, 2016). This analysis will be further discussed in the results section.

At the end of the experiment, participants were asked to report their annual income using these five responses choices: *Under 20,000, Between 21–40,000, Between 41,000–60,000, Between 61,000–80,000, Over 80,000*. They also responded to one question relating to shopping experience: *Which represents how often you shop at the major retailer X*[1]*?* Participants were given the following three response choices: *"Never," "A few times a year," "Once a month or more frequently."*

Results

If participants were able to extract an ensemble percept of product value, we would expect that their ratings of average set value would highly correlate with the retailer's value (the "predicted value"). We used a bivariate correlation to analyze participants' data, and we found significant correlations between observers' ratings of the average, or ensemble value, and the mathematical mean of the single product prices as listed by the retailer. Figure 2B shows the correlation for a single representative observer in the whole set condition (Pearson $R = 0.76$, $p < .001$). We refer to these correlations as a measure of *ensemble sensitivity*. For single subjects, the R coefficient measures ensemble sensitivity; however, the R coefficient is not normally distributed (the variance of R decreases as it approaches 1) Therefore, to measure group performance, we transformed each subjects' Pearson r values to Fisher z scores, thereby normalizing the correlation coefficient (Fisher, 1915). Then, we averaged across the 90 subjects' Fisher z scores (Average Fisher $z = 0.72$, permuted sign test, $p < .001$). The results confirmed that observers were sensitive to average price of the sets of products, even though no explicit price information was given.

If observers based their value estimate on an ensemble estimate, they should integrate stimuli from across the product set and not merely randomly sample one product from the set. To ensure participants integrated multiple products into their ensemble percept of value, we employed the subset integration measure, a method that involves presenting subsets of the whole set of objects. (Chong et al., 2008; Piazza et al., 2013; Sweeny et al., 2013, 2015; Sweeny & Whitney, 2014; Wolfe et al., 2015; Yamanashi Leib et al., 2014, 2016). On each trial, a set of 6 randomly selected consumer products was generated, with a corresponding predicted average value for that set. However, in 25 of these trials, we displayed only 4 of the 6 products to participants. In 25 trials, we displayed only 2 of the 6 products to participants, and in 25 trials, we displayed only 1 of the 6 products to participants. Trials with "subsets" of consumer products were randomly interleaved throughout the experiment. These conditions allow us to empirically simulate what participants' value estimates would be if they randomly sampled subsets of products from the whole set, instead of integrating all of the products in the set (Figure 2C & D). If participants only randomly sampled one item from the set on each trial, their ensemble value sensitivity (i.e., correlation with the predicted value) would be low and remain relatively stable, even if information (the number of products) were increasingly available to the participant (Figure 2C). In contrast, if participants integrate multiple objects into their ensemble percept, their ensemble value sensitivity should increase as more information is given to the participants (i.e., as more products are displayed; Figure 2D). Figure 2D shows the pattern of performance associated with the most extreme version of ensemble perception, with 100% of the items integrated.

Importantly, participants integrated multiple products in their ensemble valuations. Figure 2E (gray solid line) illustrates the strong subset pattern across the participants. As we revealed more information to the participants, the correlation increased (average Fisher z for 1 product = .448; Average Fisher z for 2 products = .629; Average Fisher z for 4 products = .642; Average Fisher z for 6 products = .720). A one-way, repeated measures ANOVA confirmed a main effect of set size, $F(3, 267) = 30.854$, $p < .001$, $\eta^2 = .257$, observed power = 1.00 illustrating that participants exhibited increasing ensemble sensitivity as set size increased. We also investigated the difference between performance in the 4-item subset and the whole set. Participants exhibited a significantly

[1] Name excluded for publication.

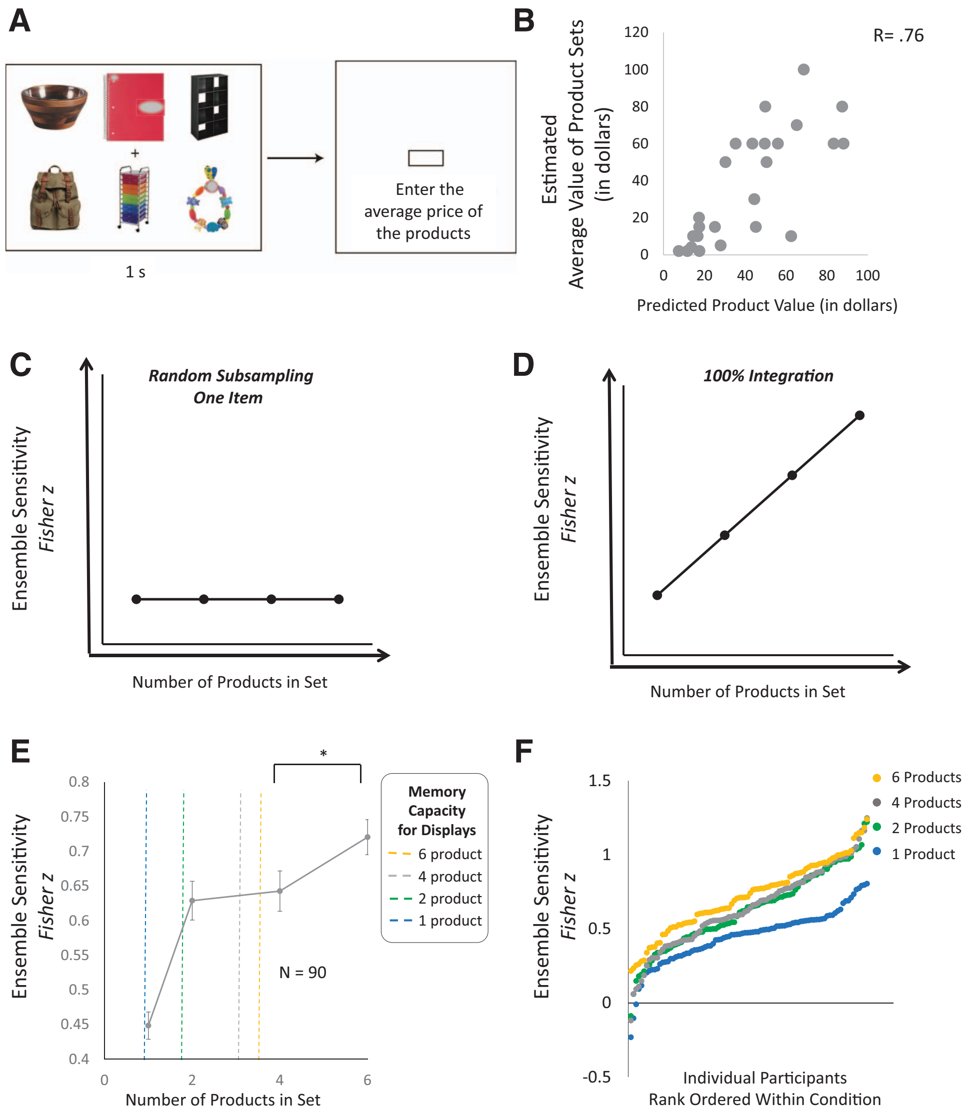

Figure 2. Experiment 1a procedure, hypothetical outcomes, and experimental results. (A) Participants viewed a set of 6 products for 1 s. No explicit price information was given. Then, participants estimated the average price of the product set. (B) A representative subject's results. The ensemble estimate of the 6-product-sets is plotted on the *y*-axis; the predicted value of the product sets, derived from retailer's listed prices, is plotted on the *x*-axis. The robust correlation suggests that this participant was able to extract ensemble value from the display of products. Group data is shown in panel E. (C) The predicted pattern of performance if observers based their judgments on one randomly sampled product from each set: Ensemble sensitivity would remain relatively constant even as more information (more products) were revealed to the participant. The *y*-axis plots the ensemble sensitivity while the *x*-axis plots the number of products displayed to the participants. (D) The predicted pattern of performance if observers integrate information from 100% of the products in the set (an extreme version of ensemble coding). The ensemble sensitivity increases as observers are given more information (as more products are revealed to the observers). (E) Empirical group data. The *y*-axis plots the magnitude of the correlation between the observers' ensemble ratings and the predicted price (i.e., ensemble sensitivity). The *x*-axis plots the number of products in the set. In the empirical data, the ensemble sensitivity increases as the number of presented products increases. The dotted lines represent participants' memory capacity measured in Experiment 2 for various set sizes. (F) Each subject's ensemble sensitivity is rank ordered within each of the different display conditions (thus, a randomly drawn vertical line at any point on the graph will likely intersect different subjects). Each dot represents one participant, but the graph is tantamount to treating all subjects as one supersubject. Different colors represent the number of products in each display condition. See the online article for the color version of this figure.

higher correlation in the whole set (Fisher z = 0.720) compared to the 4-item subset (Fisher z = 0.642), $t(89) = 2.53$, $p = .01$. Figure 2F illustrates the individual subject data for all 90 participants treated as if they are a single subject (within each subset display condition, the *x*-axis shows all 90 subjects' rank ordered ensemble sensitivities; each condition was ranked separately, for the purpose of visualization). Taken together, these results demonstrate that participants extracted an ensemble percept from the whole display, and they did not merely randomly sample one item (or even a small subset of items) from the product display. Instead, participants integrated significantly more than 4 items from the set of products.

This analysis confirms that participants did not randomly sample a small number of items from the set. Yet, is it possible that participants used something like the range or the extreme values in the product sets? To address this, we correlated each participants' ratings of the six-product ensemble with the mean of all 6 products (the original analysis, e.g., right-most data point in Figure 2E) and repeated this analysis using predicted mean prices based on the most and least expensive products in each set. Because the two distributions of predicted values are highly correlated (Figure 3A), we do not expect a large difference in the results of the two analyses. Nonetheless, there are differences in the distributions, and if participants based their ensemble judgments on the mean of 6 products, rather than the range, their ensemble sensitivity should be higher. Ensemble sensitivity was, in fact, more selective to the mean of all 6 products in the set (Fisher z = .72,) compared to mean of the highest and lowest priced products in each set (Fisher z = .66), $t(93) = 5.75$, $p < .001$, $d = .27$, observed power = .72 (Figure 3B).

At the end of the Experiment 1a, participants responded to a demographic question (5 response choices ranging from low to high income) and shopping experience question (3 response choices ranging from low to high experience shopping at the retailer; see methods). We conducted a one-way ANOVA comparing ensemble sensitivity (Fisher z on the whole set) with shopping experience and found no significant main effect. We also conducted a one-way ANOVA comparing ensemble sensitivity (Fisher z on the whole set) and income and found no significant main effect. As this was not the primary manipulation of our experiment and was merely a follow-up question, we cannot draw any strong conclusions from these results. However, broadly speaking, these results are consistent with our hypothesis that participants engaged the mechanism of ensemble perception, which does not entirely rely on experience or memory for all items in the set (Whitney & Yamanashi Leib, 2018). Experiment 3 will provide a more rigorous investigation of this question.

Experiment 1b: General Agreement of Product Value Using Independent Observers

Experiment 1a demonstrated that observers were able to extract ensemble value using the retailer's price as a baseline. In Experiment 1a, we used prices from a popular major retailer that we assumed was representative of a standard shopping experience. Nonetheless, it is customary for prices to vary slightly across various retailers in the marketplace. In Experiment 1b, we aimed to confirm generalizability of our results using a different pricing baseline. It is increasingly common for retailers to utilize some form of crowdsourcing as a significant determining factor in

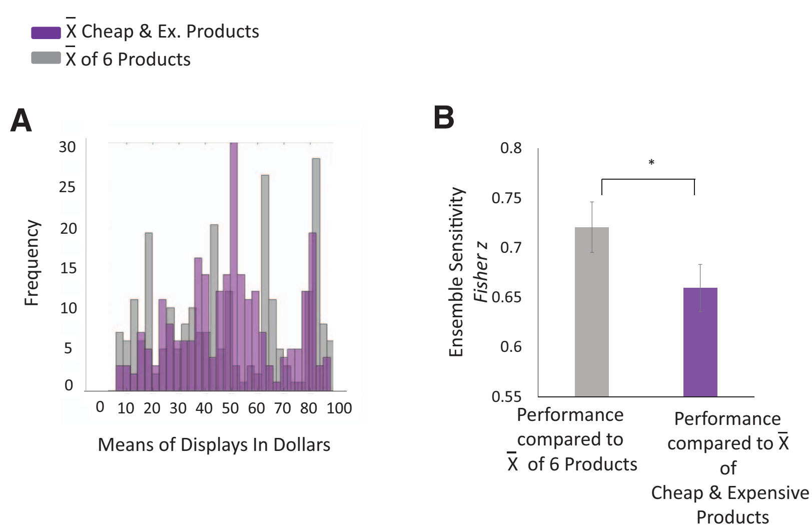

Figure 3. Experiment 1a control analysis. Alternative predicted prices for product sets. (A) Histograms of predicted mean price (in dollars) of the product sets displayed to observers. The gray color shows the predicted mean values of the 6 product sets, as was used in Figure 2B. The purple color represents an alternative predicted value, based on the average of the most and least expensive items in those sets. The histograms overlap somewhat but there are critical differences across the distribution. (B) Ensemble sensitivity to the mean of the 6-product set (gray bar) versus sensitivity recalculated relative to the mean of the most and least expensive items in the set (purple bar). Ensemble sensitivity is significantly higher to the mean price of all 6 products, confirming that participants did not base their decision on a subsample of the least and most expensive items. * $p < .001$. See the online article for the color version of this figure.

setting the price of products (Bertini & Koenigsberg, 2014). Thus, rather than picking a single competitor store to test in Experiment 1b, we used crowdsourcing, as this should yield high generalizability to a variety of stores across the marketplace. We asked independent observers on Mechanical Turk to rate the price of the products from Experiment 1a. The independent observers were shown the individual products but were not explicitly told the name of the original retailer or the listed prices. Using these independent ratings, we reexamined the ensemble product ratings from Experiment 1a and investigated whether we would observe a similar pattern of results.

Method

Participants. Experiment 1b included 100 observers recruited from Amazon Mechanical Turk (63 females, 37 male). The mean age of the participants was 35.94 ($SD = 13.50$). We asked participants to proceed with the experiments only if they had no neurological history and normal or corrected-to-normal vision. All participants were consented in accordance with the Institutional Review Board at UC Berkeley. All methods and experimental procedures were approved by the Institutional Review Board at UC Berkeley.

Stimuli and procedure. We used the same stimuli described in Experiment 1a. In Experiment 1b, participants viewed a single product for 1 s. The product was displayed centrally on a white background within a 218 × 218 pixel region. After the product disappeared, participants entered their price estimate of the product in a blank text box displayed centrally on the screen. Participants were not given a time limit to respond; they were not allowed to move to the next trial until a response was entered. There were 100 trials in all. See Figure 4A for an illustration of the trial sequence.

Results

In order to assess the reliability of the raters on single product value evaluations, we performed an intraclass correlation coefficient test, or ICC. Specifically, we used the random, two-way ICC model, measuring consistency across the average ratings. The test yielded an ICC within the excellent range, ICC = 0.819 (Cicchetti, 1994). We also investigated whether the raters agreed with the listed prices provided by the retailer. We performed a bivariate correlation and found that the raters' average estimate of product price significantly correlated with the retailer's listed price, Pearson $r = .809$, $p < .001$. Individual data (as opposed to averaged data) is also highly correlated with the retailer's listed price. For each subject, we transformed the Pearson r values to Fisher z scores and averaged across the 100 subjects, Fisher $z = 0.7619$, $p < .001$.

Finally, we compared the independent observers' ratings of single products in Experiment 1b to the observers' ensemble ratings of sets of products in Experiment 1a. We substituted the independent observers' rated prices as the predicted value on the

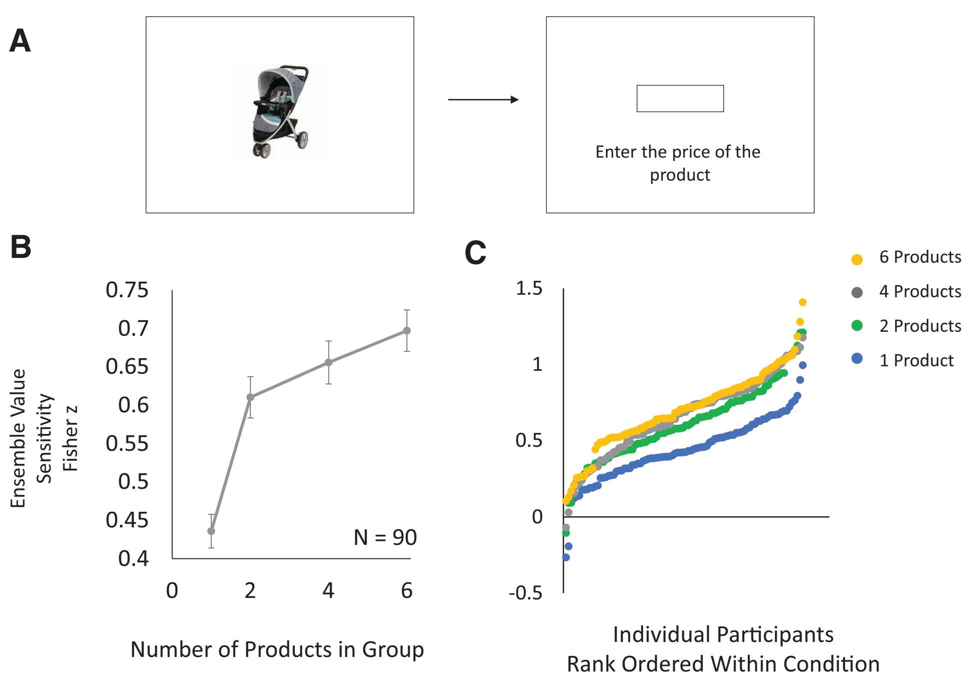

Figure 4. Example of independent observers rating procedure and results: (A) Participants viewed a single product for 1 s, then estimated its price. (B) Group data. The *y*-axis plots the magnitude of the correlation between the observers' ensemble ratings and the predicted price using independent observers (i.e., ensemble sensitivity). The *x*-axis plots the number of products in the set. The ensemble sensitivity increases as the size of the set increases, indicating that participants are integrating multiple products into their ensemble percept. (C) Individual data. The *y*-axis plots the ensemble sensitivity. The *x*-axis plots subjects' data in ranked order of performance within each condition. Each dot represents a subject; the different colors represent the number of products in each display condition. See the online article for the color version of this figure.

x-axis (cf., Figure 2B). Individual subjects' ratings of average set price (Experiment 1a) correlated highly with the independent observer's set price ratings derived from single product ratings in Experiment 1b, Fisher z = 0.71, permuted sign test, $p < .001$. This result is unsurprising because initial analyses demonstrated that observers' ratings in Experiment 1a correlated highly with retailer's price. This merely serves as further confirmation that observers largely agreed about the abstracted value of consumer products. We also observed the expected subset integration pattern. In Experiment 1a, we included displays with different numbers of subsets of products (showing 4 products out of 6, 2 products out of 6, and 1 product out of 6). These conditions allow us to empirically simulate what participants' ensemble estimates would be if they randomly sampled subsets of products from the whole set, instead of integrating all of the products in the set. We measured participants' ensemble sensitivity to varying subsets relative to the mean of the whole set: Subset with 1 product, Fisher z = .434; Subset with 2 products, Fisher z = .610; Subset with 4 products, Fisher z = .656; Whole set with 6 products, Fisher z = .697. We observed the expected monotonic increase in ensemble sensitivity as more products were revealed (Figure 4B). A One-Way repeated measure ANOVA confirmed a main effect of set size, $F(3, 267) = 26.751$, $p < .001$, $\eta^2 = .231$, observed power = 1.00. This confirms that participants integrated multiple items into their ensemble judgment of product value using ratings of independent observers. For the purpose of visualization, we rank ordered participants' ensemble sensitivity within the different set size conditions (Figure 4C).

Experiment 2: Explicit Memory for Individual Set Members

Several previous studies have demonstrated that single item recollection is not a prerequisite for ensemble perception in orientation, size, and facial expression (Ariely, 2001; Fischer & Whitney, 2011; Haberman & Whitney, 2011; Parkes et al., 2001). Experiment 3 investigated whether the same is true for consumer product ensemble judgments. One possibility is that participants extracted the average value of consumer products by explicitly remembering every item in the set. Another alternative is that the ensemble judgment occurs independently of explicit memory for every product. Specifically, we explored whether the display duration of Experiment 1 was sufficiently long for participants to explicitly recall each member in the set.

Method

Participants. Experiment 2 included 85 observers recruited from Amazon Mechanical Turk (37 females, 48 males). The mean age of the participants was 36.25 ($SD = 11.80$). We asked participants to proceed with the experiments only if they had no neurological history and normal or corrected-to-normal vision. All participants were consented in accordance with the Institutional Review Board at UC Berkeley. All methods and experimental procedures were approved by the Institutional Review Board at UC Berkeley.

Stimuli and procedure. Experiment 2 used the same stimuli from Experiments 1a & 1b. In each trial, observers viewed 6 randomly chosen stimuli for 1 s (identical to the display duration in Experiment 1). After the 6 stimuli disappeared, observers performed a membership identity task. Two stimuli appeared on the screen. One stimulus was a member of the previously displayed product set and one was a lure (randomly drawn from the entire stimulus array). Participants were asked to pick the member of the set using a button press. In addition to sets of 6 products (as in Experiment 1), we also displayed trials with 1, 2, and 4 products in each set. Observers participated in 100 trials total (25 in each set size).

Results

To determine how many products observers were able to recall in each display, we used the following calculation: $MC = I \times P$, where MC represents working memory capacity, I represents the number of items in the set, and P represents the proportion of items remembered. Specifically, the formula was: $\frac{(trialscorrect - trialscorrectbychance) \times 2}{totalnumberoftrials}$. Because we employed a 2AFC experimental design, we define *trials correct by chance* as 50% of the available trials. Participants remembered approximately 1 product in the 1-product trials, 2 products in the 2-product trials, and less than 4 products in the 4-product and 6-product trials. (1-product trials: $M = 0.90$ product, SEM = 0.02; 2-product trials: $M = 1.74$ products, SEM = 0.056; 4-product trials: $M = 3.06$ products, SEM = 0.096; 6-product trials: $M = 3.54$ products SEM = 0.171). In contrast, participants integrated significantly more than four products and up to 6 items (See Figure 2E). This finding suggests that ensemble judgments of consumer products are based on more than an explicit memory of each exemplar. The experience and income questions reported in Experiment 1a were also asked at the end of this experiment. We repeated the one-way ANOVAs comparing shopping experience and memory and income and memory. We found no significant main effects.

Experiment 3: Products From the Same Category: Diverse Quality and Prices

The prior experiments demonstrate that observers are sensitive the ensemble value of consumer products. Furthermore, ensemble perception does not depend on memory for individual products in the display. The original stimulus array contained consumer products from diverse product categories and included only a few products that shared the same category. Many shopping experiences include comparisons of products from the same category that vary in quality and/or price. Therefore, to further increase the generalizability of our results, we reran the experiment using a new array of stimuli that incorporated more products drawn from within the *same* category.

Method

Participants. Experiment 3 included 90 observers recruited from Amazon Mechanical Turk (33 females, 57 males). The mean age of the participants was 36.65 ($SD = 10.57$). We asked participants to proceed with the experiments only if they had no neurological history and normal or corrected-to-normal vision. All participants were consented in accordance with the Institutional Review Board at UC Berkeley. All methods and experimental procedures were approved by the Institutional Review Board at UC Berkeley.

Stimuli and procedure. In Experiment 3, we created a new array of stimuli that explicitly contained product pairs from the 26-product category that differed in price and quality (Figure 5A). For example, the set contained a relatively expensive brand of coffee compared to a more common brand of coffee. A comprehensive description of the set of 52 new stimuli is provided in the online supplemental materials. We generated sets of stimuli by randomly drawing 6 products without replacement from the 52 products. Observers viewed 30 trials of 6 products and 30 trials in each of the remaining subsets of 4, 2, and 1 consumer products. Subset size was randomly interleaved across trials. Participants were not required to fixate in the center of the screen. All other aspects of the experiment were identical to Experiment 1a.

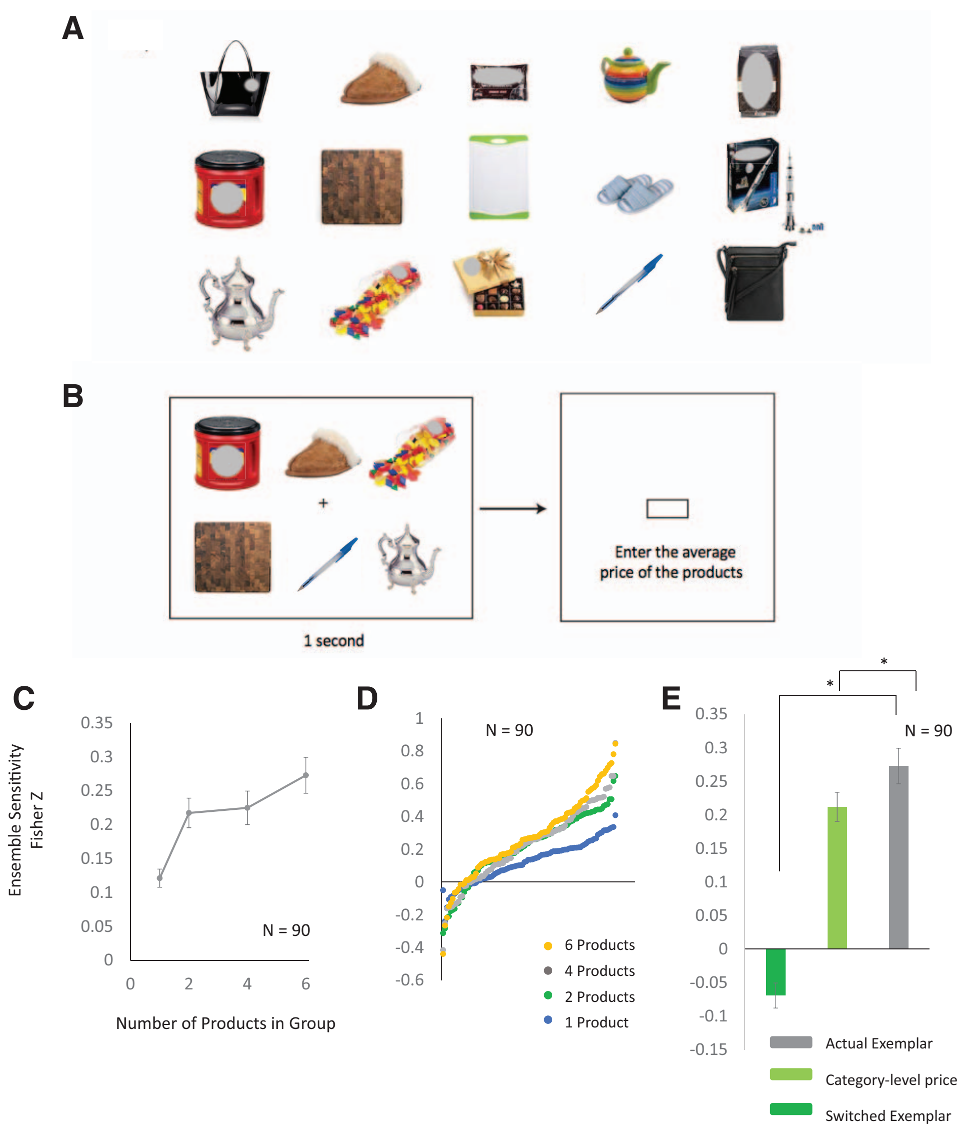

Figure 5. Example of Experiment 3 stimulus array, procedure, and results: (A) The new stimulus array consisted of 52 products that were drawn from 26 categories. Products from within a given category differed in price and/or quality. (B) Observers viewed 6 products randomly selected from the stimulus array for 1 s, and then estimated the average price of the products. (C) Experiment 3 Results. Ensemble sensitivity (*y*-axis) increases with increasing set size (*x*-axis), indicating that participants integrated multiple products into their ensemble percept. (D) Rank-ordered ensemble sensitivity within each set size. Each dot represents a subject, and the graph represents a single supersubject. (E) Participants' responses modeled by three different predictors: The price of the actual exemplars displayed (using predicted product value; gray bar); the price of an exemplar from the same category (the pseudopredicted value; dark green bar); and the average price of the exemplars within each category (light green bar). Participants exhibit ensemble sensitivity to the specific exemplars displayed. * $p <$ .001. See the online article for the color version of this figure.

Results

As in all the prior experiments, if participants were able to extract an ensemble percept of product value, we would expect that their ratings of average set value would correlate with the retailer's value. We used a bivariate correlation to analyze participants' data, and we found significant correlations between observers' rating of average, or ensemble value, and the mathematical mean of the single product ratings as listed by the retailer. We normalized individual subjects' correlation coefficients using Fisher's Z transformation and averaged across the group of 90 participants. The results confirm that the group of participants exhibited ensemble sensitivity to the retailer's listed value, average Fisher z = .273, permuted sign test, $p < .001$.

Ensemble sensitivity to varying subsets was as follows: Subset with 1 product, Fisher z = .112; Subgroup with 2 products, Fisher z = .217; Subset with 4 products, Fisher z = .225; Whole Set with 6 products, Fisher z = .273. A One-Way repeated measure ANOVA confirmed a main effect of set size, $F(3, 267) = 14.359$, $p < .001$, $\eta^2 = .139$, observed power = 1.00. This confirms that participants integrated multiple items into their ensemble judgment of product value (Figure 5C and 5D).

While all results are statistically significant, the overall ensemble sensitivity was lower compared to Experiment 1. This is expected because the variance of the displays was greater compared to the prior experiments (mean variance of displays in Experiment 1a = 26.01; mean variance of displays in Experiment 1b = 33.87; mean variance in Experiment 3 = 44.68). It is well-established that high variance negatively impacts ensemble sensitivity (Dakin, 2001; Fouriezos, Rubenfeld, & Capstick, 2008; Haberman et al., 2015; Im & Halberda, 2013; Marchant et al., 2013; Morgan, Chubb, & Solomon, 2008). The high variance was necessary to conduct the two forthcoming analyses that verify participants successfully discriminated exemplars within product categories.

To ensure that participants discriminated between the quality of the different exemplars within each product category, we ran two additional analyses. First, we reran the analysis of the 6-products sets using an alternate pseudopredicted value for the product sets. Specifically, we switched the prices of products *within* category. For example, if a participant viewed a display containing a silver teapot, a plastic office pen, a butcher-block wood cutting board, name-brand slippers, and off-brand toy blocks, we replaced the prices with those of a ceramic teapot, a luxury pen, a plastic cutting board, off-brand slippers, and name-brand toy blocks. We then correlated participants' responses to the original display with the pseudopredicted values.

If participants incorporated exemplars of diverse quality into their ensemble percept (rather than randomly interchanging exemplar prices within a product category), we should observe that ensemble sensitivity decreases using pseudopredicted product value. We find exactly that. Participants' ensemble sensitivity dramatically decreased using the within-category swapped prices, Fisher z actual exemplars = .273; Fisher z switched exemplars = −.069. A paired t test confirms the difference is significant, $t(89) = 8.54$, $p < .001$, $d = 90$, observed power = 1.00. This analysis indicates that participants did not interchange the prices between different exemplars and also confirms that we used stimuli with sufficient price differences to rule out this plausible strategy (Figure 5E).

One additional possibility is that participants did not interchange the prices between exemplars, but rather relied on a category-level representation that typifies the price for all products within that category. To ensure participants did not rely on this strategy, we averaged the prices of the products *within* each product category. For example, our stimuli array contained a wood, butcher-block cutting board ($47.59) and a plastic cutting board ($14.80). The average price of these products was: $31.19. We repeated this averaging process within each category across the stimuli set to estimate category-level prices. Finally, we reran the analysis comparing participants' original responses to the category-level price of the products. If participants relied on broad category-level or prototype representations of value, we should observe that ensemble sensitivity is similar for both the original analysis and the new analysis (using general category-level prices as the predictive value). Instead, if participants discriminate between specific exemplars, ensemble sensitivity will decrease using the category-level analysis. As expected, ensemble sensitivity decreased when general category-level prices were substituted for specific exemplar prices. A paired t test confirms the difference is significant, $t(89) = 3.52$, $p < .001$, $d = .371$, observed power = .94. This establishes that participants discriminated the quality of exemplar products within each category, and did not rely on a category-level, generic, or prototype judgments.

Discussion

To investigate whether ensemble perception can operate across abstracted impressions, we asked observers to report the ensemble value from diverse sets of consumer products. We found that observers successfully report the ensemble value—and they do so with limited exposure time and without explicit memory of each individual product. Because the price of the products was not explicitly presented, observers integrated abstracted impressions to successfully complete the ensemble task. This is strong evidence that ensemble perception allows us to rapidly derive abstract gist information from our visual surroundings. These findings may also have practical relevance for consumer marketing. Observers' assessments of product value were broadly consistent with retailers' prices and with each other. This suggests that ensemble percepts of value are shared among observers and accessible to a broad population of consumers.

Ensemble perception operates over a broad range of visual (Alvarez, 2011; Hochstein, Pavlovskaya, Bonneh, & Soroker, 2015; Whitney & Yamanashi Leib, 2018) and auditory attributes (Albrecht, Scholl, & Chun, 2012; McDermott, Schemitsch, & Simoncelli, 2013; Piazza et al., 2013), including high level visual information like facial expression (Haberman & Whitney, 2007, 2009), animacy (Yamanashi Leib et al., 2016), and attractiveness (Post, Haberman, Iwaki, & Whitney, 2012). However, there is no clear consensus regarding whether ensemble perception operates across abstracted visual impressions. Post et al., displayed photographs of crowds, and asked observers to report the average attractiveness of the individuals in the crowds. Participants successfully perceived the ensemble attractiveness of the crowds, suggesting a minimal level of perceptual abstraction. However, in this case, ensemble attractiveness could also be driven by low-level visual features, such as symmetry (Perrett et al., 1999; Rhodes, Sumich, & Byatt, 1999). More recently, Yamanashi Leib, et al. showed participants brief displays of household objects, animals, and people and asked observers to rate the overall animacy of the display. They reported high sensitivity to the ensemble animacy of the displays (Yamanashi Leib, et al., 2016). Again, this might suggest that ensemble perception operates beyond a basic visual feature analysis. However, perceptions of animacy are also associated with low-level features, such as perceived curvature of the stimuli (Long, Störmer, & Alvarez, 2017). The current findings represent strong evidence that ensemble coding mechanisms can integrate abstracted impressions of value into a single, unitary percept.

Perceived value is a complex construct. It reflects many different dimensions including product quality, product performance, emotional significance, social usefulness, and price, among other factors (Sweeney & Soutar, 2001). Our experiments show that observers can ensemble code the utility of groups of objects as evidenced by their ability to translate product groups into ensemble price estimates. Reported price is simply a convenient corollary for value, but which dimensions are used, and whether individual objects are transformed into other intermediate dimensions before the averaging occurs, are open questions. For example, observers might transform the visual images of products into a continuous representation of emotional significance, or functional utility, or dollar value, or a quality metric, or something else, before integrating this information into a reported ensemble price. In fact, it is likely that subjects actually use a combination of these dimensions to assess the value of any given object (Sweeney & Soutar, 2001). Likewise, integrating across any or all of these dimensions to estimate the perceived value of a crowd of objects, as we found here, counts as an ensemble percept of value.

The question remains what neural mechanisms might underlie this unique phenomenon. Ensemble information is likely represented at many stages of the visual processing stream (Haberman, Brady, & Alvarez, 2015; Hubert-Wallander et al., 2015; Whitney & Yamanashi Leib, 2018). For example, low-level feature ensembles are represented in visual cortex (Britten, Shadlen, Newsome, & Movshon, 1992) and more recent research has found neural populations that carry ensemble shape and face information (Cant & Xu, 2017; Im et al., 2017). However, whether these same mechanisms could be used for representing product value or utility remains unclear. Future research should therefore investigate the neural mechanisms that contribute to abstracted ensemble perception.

Our results confirm that ensemble value can be successfully extracted when the individual products cannot be explicitly recalled. This finding highlights the efficiency of ensemble coding consumer product value, even under conditions where other visual processes, such as visual working memory, are constrained. Some prior research has shown that ensemble representations can be achieved without recalling every item (e.g., circle size, orientation, faces, etc. Alvarez, 2011; Ariely, 2001; Fischer & Whitney, 2011; Haberman & Whitney, 2010; Hochstein et al., 2015; Parkes et al., 2001). Our research demonstrates that abstract values assigned to objects can be achieved rapidly—without reliance on a maintained representation of each product. Although participants cannot remember each item, they are nonetheless able to integrate most or all of the items into an accurate ensemble valuation of the whole product set.

Observers need not integrate every member of a set to perceive and report a summary statistic (Whitney & Yamanashi Leib, 2018). Diverse patterns of integration between subset sizes can also be observed, potentially suggesting different strategies are used by the visual system when set size varies (Marchant et al., 2013). Nevertheless, our results suggest that most, if not all, of the objects were integrated into an ensemble estimate. To empirically demonstrate this, we used the subset integration measure, a standard method used in ensemble perception paradigms (Chong et al., 2008; Piazza et al., 2013; Sweeny & Whitney, 2014; Sweeny et al., 2013, 2015; Wolfe et al., 2015; Yamanashi Leib et al., 2014, 2016). In addition, this and prior ensemble studies also demonstrate that nonrandom sampling (e.g., sampling the extreme values in the set) cannot account for performance in ensemble tasks (Ariely, 2001; Chong & Treisman, 2003; Fischer & Whitney, 2011; Haberman & Whitney, 2007, 2009; Robitaille & Harris, 2011; Wolfe et al., 2015). The integration efficiency—the number of integrated objects in a summary statistical percept—can vary, though a rough rule of thumb is that approximately the square root of the number of objects are integrated in many different types of ensemble percepts (Whitney & Yamanashi Leib, 2018). Our results here fall on the high side of that range, but they are consistent with many other studies on ensemble perception with similar exposure durations (Gorea, Belkoura, & Solomon, 2014; Yamanashi Leib et al., 2014, 2016).

A great deal is known about consumer perception and attitudes toward individual objects, but much less effort has been devoted to understanding the initial impressions that consumers form about object displays. In these initial impressions, collections of objects are the rule, not the exception. Our results fill a unique gap, revealing a fast and efficient mechanism for representing the summary statistical value of product sets. Our results raise the intriguing possibility that value-assessments at a glance, and the underlying ensemble perception mechanisms that support these judgments, may shape early stage evaluations of product displays. Moreover, these findings confirm that ensemble perception, a visual mechanism that provides short-cuts in cluttered visual environments, can integrate abstracted visual information.

References

Albrecht, A. R., Scholl, B. J., & Chun, M. M. (2012). Perceptual averaging by eye and ear: Computing summary statistics from multimodal stimuli. *Attention, Perception, & Psychophysics, 74,* 810–815. http://dx.doi.org/10.3758/s13414-012-0293-0

Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences, 15,* 122–131. http://dx.doi.org/10.1016/j.tics.2011.01.003

Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science, 12,* 157–162. http://dx.doi.org/10.1111/1467-9280.00327

Bertini, M., & Koenigsberg, O. (2014). When customers help set prices. *MIT Sloan Management Review, 55,* 57.

Bishop, W. R., Jr. (1984). Competitive intelligence. *Progressive Grocer, 63,* 19–20.

Bonner, P. G., & Nelson, R. (1985). Product attributes and perceived quality: Foods. In J. Jacoby & J. Olson (Eds.), *Perceived quality* (pp. 65–79). Lexington, MA: Lexington Books.

Britten, K. H., Shadlen, M. N., Newsome, W. T., & Movshon, J. A. (1992). The analysis of visual motion: A comparison of neuronal and psychophysical performance. *The Journal of Neuroscience, 12,* 4745–4765. http://dx.doi.org/10.1523/JNEUROSCI.12-12-04745.1992

Cant, J. S., & Xu, Y. (2017). The contribution of object shape and surface properties to object ensemble representation in anterior-medial ventral visual cortex. *Journal of Cognitive Neuroscience, 29,* 398–412. http://dx.doi.org/10.1162/jocn_a_01050

Chong, S. C., Joo, S. J., Emmanouil, T. A., & Treisman, A. (2008). Statistical processing: Not so implausible after all. *Perception & Psychophysics, 70,* 1327–1334. http://dx.doi.org/10.3758/PP.70.7.1327

Chong, S. C., & Treisman, A. (2003). Representation of statistical properties. *Vision Research, 43,* 393–404. http://dx.doi.org/10.1016/S0042-6989(02)00596-5

Chong, S. C., & Treisman, A. (2005). Statistical processing: Computing the average size in perceptual groups. *Vision Research, 45,* 891–900. http://dx.doi.org/10.1016/j.visres.2004.10.004

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6,* 284–290. http://dx.doi.org/10.1037/1040-3590.6.4.284

Dakin, S. C. (2001). Information limit on the spatial integration of local orientation signals. *Journal of the Optical Society of America, A, Optics, Image Science & Vision, 18,* 1016–1026. http://dx.doi.org/10.1364/JOSAA.18.001016

Demeyere, N., Rzeskiewicz, A., Humphreys, K. A., & Humphreys, G. W. (2008). Automatic statistical processing of visual properties in simultanagnosia. *Neuropsychologia, 46,* 2861–2864. http://dx.doi.org/10.1016/j.neuropsychologia.2008.05.014

Doyle, M. (1984). New ways of measuring value. Progressive grocer-value, *Executive Report,* 15–19.

Elias, E., Dyer, M., & Sweeny, T. D. (2017). Ensemble perception of dynamic emotional groups. *Psychological Science, 28,* 193–203. http://dx.doi.org/10.1177/0956797616678188

Fischer, J., & Whitney, D. (2011). Object-level visual information gets through the bottleneck of crowding. *Journal of Neurophysiology, 106,* 1389–1398. http://dx.doi.org/10.1152/jn.00904.2010

Fisher, R. A. (1915). Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika, 10,* 507–521. http://dx.doi.org/10.2307/2331838

Florey, J., Clifford, C. W., Dakin, S., & Mareschal, I. (2016). Spatial limitations in averaging social cues. *Scientific Reports, 6,* 32210. http://dx.doi.org/10.1038/srep32210

Florey, J., Dakin, S. C., & Mareschal, I. (2017). Comparing averaging limits for social cues over space and time. *Journal of Vision, 17*(9), 17. http://dx.doi.org/10.1167/17.9.17

Fouriezos, G., Rubenfeld, S., & Capstick, G. (2008). Visual statistical decisions. *Perception & Psychophysics, 70,* 456–464. http://dx.doi.org/10.3758/PP.70.3.456

Goodale, B. M., Alt, N. P., Lick, D. J., & Johnson, K. L. (2018). Groups at a glance: Perceivers infer social belonging in a group based on perceptual summaries of sex ratio. *Journal of Experimental Psychology: General, 147,* 1660–1676. http://dx.doi.org/10.1037/xge0000450

Gorea, A., Belkoura, S., & Solomon, J. A. (2014). Summary statistics for size over space and time. *Journal of Vision, 14*(9), 22. http://dx.doi.org/10.1167/14.9.22

Haberman, J., Brady, T. F., & Alvarez, G. A. (2015). Individual differences in ensemble perception reveal multiple, independent levels of ensemble representation. *Journal of Experimental Psychology: General, 144,* 432–446. http://dx.doi.org/10.1037/xge0000053

Haberman, J., & Whitney, D. (2007). Rapid extraction of mean emotion and gender from sets of faces. *Current Biology, 17,* R751–R753. http://dx.doi.org/10.1016/j.cub.2007.06.039

Haberman, J., & Whitney, D. (2009). Seeing the mean: Ensemble coding for sets of faces. *Journal of Experimental Psychology: Human Perception and Performance, 35,* 718–734. http://dx.doi.org/10.1037/a0013899

Haberman, J., & Whitney, D. (2010). The visual system discounts emotional deviants when extracting average expression. *Attention, Perception, & Psychophysics, 72,* 1825–1838. http://dx.doi.org/10.3758/APP.72.7.1825

Haberman, J., & Whitney, D. (2011). Efficient summary statistical representation when change localization fails. *Psychonomic Bulletin & Review, 18,* 855–859. http://dx.doi.org/10.3758/s13423-011-0125-6

Hochstein, S., Pavlovskaya, M., Bonneh, Y. S., & Soroker, N. (2015). Global statistics are not neglected. *Journal of Vision, 15*(4), 7. http://dx.doi.org/10.1167/15.4.7

Hubert-Wallander, B., & Boynton, G. M. (2015). Not all summary statistics are made equal: Evidence from extracting summaries across time. *Journal of Vision, 15*(4), 5.

Im, H. Y., Albohn, D. N., Steiner, T. G., Cushing, C. A., Adams, R. B., Jr., & Kveraga, K. (2017). Differential hemispheric and visual stream contributions to ensemble coding of crowd emotion. *Nature Human Behaviour, 1,* 828–842. http://dx.doi.org/10.1038/s41562-017-0225-z

Im, H. Y., & Halberda, J. (2013). The effects of sampling and internal noise on the representation of ensemble average size. *Attention, Perception, & Psychophysics, 75,* 278–286. http://dx.doi.org/10.3758/s13414-012-0399-4

Jacoby, J., & Olson, J. C. (1985). *Perceived quality*. Lexington, MA: Lexington Books.

Laird, D. A. (1932). How the consumer estimates quality by subconscious sensory impressions. *Journal of Applied Psychology, 16,* 241–246. http://dx.doi.org/10.1037/h0074816

Long, B., Störmer, V. S., & Alvarez, G. A. (2017). Mid-level perceptual features contain early cues to animacy. *Journal of Vision, 17*(6), 20. http://dx.doi.org/10.1167/17.6.20

Marchant, A. P., & de Fockert, J. W. (2009). Priming by the mean representation of a set. *The Quarterly Journal of Experimental Psychology, 62,* 1889–1895. http://dx.doi.org/10.1080/17470210902871045

Marchant, A. P., Simons, D. J., & de Fockert, J. W. (2013). Ensemble representations: Effects of set size and item heterogeneity on average size perception. *Acta Psychologica, 142,* 245–250. http://dx.doi.org/10.1016/j.actpsy.2012.11.002

McDermott, J. H., Schemitsch, M., & Simoncelli, E. P. (2013). Summary statistics in auditory perception. *Nature Neuroscience, 16,* 493–498. http://dx.doi.org/10.1038/nn.3347

Morgan, M., Chubb, C., & Solomon, J. A. (2008). A "dipper" function for texture discrimination based on orientation variance. *Journal of Vision, 8*(11), 9. http://dx.doi.org/10.1167/8.11.9

Olshavsky, R. W. (1985). Perceived quality in consumer decision making: An integrated theoretical perspective. *Perceived Quality, 4,* 3–29.

Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience, 4,* 739–744. http://dx.doi.org/10.1038/89532

Perrett, D. I., Burt, D. M., Penton-Voak, I. S., Lee, K. J., Rowland, D. A., & Edwards, R. (1999). Symmetry and human facial attractiveness. *Evolution and Human Behavior, 20,* 295–307. http://dx.doi.org/10.1016/S1090-5138(99)00014-8

Piazza, E. A., Sweeny, T. D., Wessel, D., Silver, M. A., & Whitney, D. (2013). Humans use summary statistics to perceive auditory sequences. *Psychological Science, 24,* 1389–1397. http://dx.doi.org/10.1177/0956797612473759

Post, R. B., Haberman, J., Iwaki, L., & Whitney, D. (2012). The frozen face effect: Why static photographs may not do you justice. *Frontiers in Psychology, 3,* 22. http://dx.doi.org/10.3389/fpsyg.2012.00022

Rhodes, G., Sumich, A., & Byatt, G. (1999). Are average facial configurations attractive only because of their symmetry? *Psychological Science, 10,* 52–58. http://dx.doi.org/10.1111/1467-9280.00106

Robitaille, N., & Harris, I. M. (2011). When more is less: Extraction of summary statistics benefits from larger sets. *Journal of Vision, 11*(12), 18. http://dx.doi.org/10.1167/11.12.18

Sawyer, A. G., & Dickson, P. R. (1984). Psychological perspectives on consumer response to sales promotion. Research on sales promotion: Collected papers, 1–21.

Schechter, L. (1984). A normative conception of value. *Progressive Grocer: Executive Report, 2,* 12–14.

Sweeney, J. C., & Soutar, G. N. (2001). Consumer perceived value: The development of a multiple item scale. *Journal of Retailing, 77,* 203–220. http://dx.doi.org/10.1016/S0022-4359(01)00041-0

Sweeny, T. D., Haroz, S., & Whitney, D. (2013). Perceiving group behavior: Sensitive ensemble coding mechanisms for biological motion of human crowds. *Journal of Experimental Psychology: Human Perception and Performance, 39,* 329–337. http://dx.doi.org/10.1037/a0028712

Sweeny, T. D., & Whitney, D. (2014). Perceiving crowd attention: Ensemble perception of a crowd's gaze. *Psychological Science, 25,* 1903–1913. http://dx.doi.org/10.1177/0956797614544510

Sweeny, T. D., Wurnitsch, N., Gopnik, A., & Whitney, D. (2015). Ensemble perception of size in 4–5-year-old children. *Developmental Science, 18,* 556–568. http://dx.doi.org/10.1111/desc.12239

van Osch, Y., Blanken, I., Meijs, M. H. J., & van Wolferen, J. (2015). A group's physical attractiveness is greater than the average attractiveness of its members: The group attractiveness effect. *Personality and Social Psychology Bulletin, 41,* 559–574. http://dx.doi.org/10.1177/0146167215572799

Walker, D., & Vul, E. (2014). Hierarchical encoding makes individuals in a group seem more attractive. *Psychological Science, 25,* 230–235. http://dx.doi.org/10.1177/0956797613497969

Watamaniuk, S. N., & Duchon, A. (1992). The human visual system averages speed information. *Vision Research, 32,* 931–941. http://dx.doi.org/10.1016/0042-6989(92)90036-I

Watamaniuk, S. N., & McKee, S. P. (1998). Simultaneous encoding of direction at a local and global scale. *Perception & Psychophysics, 60,* 191–200. http://dx.doi.org/10.3758/BF03206028

Watamaniuk, S. N., Sekuler, R., & Williams, D. W. (1989). Direction perception in complex dynamic displays: The integration of direction information. *Vision Research, 29,* 47–59. http://dx.doi.org/10.1016/0042-6989(89)90173-9

Webster, J., Kay, P., & Webster, M. A. (2014). Perceiving the average hue of color arrays. *Journal of the Optical Society of America A, Optics, Image Science, and Vision, 31*(4), A283–A292. http://dx.doi.org/10.1364/JOSAA.31.00A283

Whitney, D., & Yamanashi Leib, A. (2018). Ensemble Perception. *Annual Review of Psychology, 69,* 105–129. http://dx.doi.org/10.1146/annurev-psych-010416-044232

Wolfe, B. A., Kosovicheva, A. A., Leib, A. Y., Wood, K., & Whitney, D. (2015). Foveal input is not required for perception of crowd facial expression. *Journal of Vision, 15*(4), 11. http://dx.doi.org/10.1167/15.4.11

Yamanashi Leib, A., Fischer, J., Liu, Y., Qiu, S., Robertson, L., & Whitney, D. (2014). Ensemble crowd perception: A viewpoint-invariant mechanism to represent average crowd identity. *Journal of Vision, 14*(8), 26. http://dx.doi.org/10.1167/14.8.26

Yamanashi Leib, A., Kosovicheva, A., & Whitney, D. (2016). Fast ensemble representations for abstract visual impressions. *Nature Communications, 7,* 13186. http://dx.doi.org/10.1038/ncomms13186

Zeithaml, V. A. (1988). Consumer perceptions of price, quality, and value: A means-end model and synthesis of evidence. *Journal of Marketing, 52,* 2–22. http://dx.doi.org/10.1177/002224298805200302

Received October 30, 2018
Revision received November 9, 2019
Accepted January 3, 2020 ■