

# Overestimating the Valuations and Preferences of Others

Minah H. Jung  
New York University

Alice Moon  
University of Pennsylvania

Leif D. Nelson  
University of California, Berkeley

People often make judgments about their own and others' valuations and preferences. Across 12 studies ( $N = 18,818$ ), we find a robust bias in these judgments such that people overestimate the valuations and preferences of others. This overestimation arises because, when making predictions about others, people rely on their intuitive core representation of the experience (e.g., is the experience generally positive?) in lieu of a more complex representation that might also include countervailing aspects (e.g., is any of the experience negative?). We first demonstrate that the overestimation bias is pervasive for a wide range of positive (Studies 1–5) and negative experiences (Study 6). Furthermore, the bias is not merely an artifact of how preferences are measured (Study 7). Consistent with judgments based on core representations, the bias significantly reduces when the core representation is uniformly positive (Studies 8A–8B). Such judgments lead to a paradox in how people see others trade off between valuation and utility (Studies 9A–9B). Specifically, relative to themselves, people believe that an identically paying other will get more enjoyment from the same experience, but paradoxically, that an identically enjoying other will pay more for the same experience. Finally, consistent with a core representation explanation, explicitly prompting people to consider the entire distribution of others' preferences significantly reduced or eliminated the bias (Study 10). These findings suggest that social judgments of others' preferences are not only largely biased, but they also ignore how others make trade-offs between evaluative metrics.

**Keywords:** overestimation bias, comparative judgments, valuation, preferences, paradox


**Supplemental materials:** <http://dx.doi.org/10.1037/xge0000700.supp>

Even the most enjoyable experiences can have shortcomings. A first trip to Rio De Janeiro may include a panoramic view from atop Sugarloaf Mountain, a tour of the Selarón Steps, and some relaxing strolls along Ipanema Beach. Nevertheless, one could also anticipate the inevitable complications that come from idiosyncratic personal preferences. Perhaps the traveler is afraid of heights, does not care much for artistic sites, or simply dislikes

strolling in general. Accordingly, although the traveler might anticipate their experience to be positive, they might not anticipate it to be uniformly so. Now consider the same person asked to anticipate the experience of a fellow traveler. The positives are shared—beautiful views are generally beautiful—but the shortcomings might not be—not everyone is afraid of heights. Forecasting our personal experience is enriched by introspection into personal preferences, but forecasting the experience of someone else will be more closely tied to the primary representation of the experience in question. A vacation to Rio De Janeiro sounds largely pleasant. So even though someone might have their own personal misgivings, they might still believe that someone else will probably enjoy everything about it.

People frequently need to make judgments of valuation similar to that of a vacation in Rio De Janeiro. But when forming these judgments both for themselves and for others, people nearly always have imperfect knowledge. Because knowledge is imperfect, error is inevitable. Moreover, because knowledge asymmetrically favors the self (Krueger, 2000), those errors can also produce bias. In this article, we investigate how people forecast valuation for self and for others and document a pervasive, robust, and irrational tendency to believe that other people evaluate things more intensely. Predicted experience intensity, we theorize, is not a result of seeing amplification in the experience of others, but rather in the uniformity of it. People predict that someone else will rate their

This article was published Online First November 21, 2019.

 Minah H. Jung, Department of Marketing, Leonard N. Stern School of Business, New York University; Alice Moon, Department of Operations, Information, and Decisions, The Wharton School, University of Pennsylvania; Leif D. Nelson, Haas School of Business, University of California, Berkeley.

Minah H. Jung and Alice Moon contributed equally to this research and are joint first authors. This research was supported in part by the Fetzer Franklin Fund of the John E. Fetzer Memorial Trust, the Wharton Dean's Research Fund, and the Center for Global Economy and Business Faculty Grant, Stern School of Business, New York University. All data, R code, and pre-registrations for all studies reported in the main text and the additional online data are available at <https://osf.io/b8253/>.

Correspondence concerning this article should be addressed to Minah H. Jung, Department of Marketing, Leonard N. Stern School of Business, New York University, 40 West 4th Street, Tisch Hall 912, New York, NY 10012. E-mail: [minah.jung@stern.nyu.edu](mailto:minah.jung@stern.nyu.edu)

vacation to Rio De Janeiro more intensely not because the positives are more amplified, but because the negatives are not sufficiently considered in that predicted rating.

Research in social judgment has documented a wide variety of errors and biases in the estimation of other people's preferences. Perhaps the most canonical insight has been the documented tendency for people to assume that other people have similar preferences to themselves (Krueger, 1998). For example, someone who would pay a parking ticket by mail, rather than contest it at the courthouse, will give a higher estimate of how many other people will pay the fine than would those who personally prefer to contest the ticket (Ross, Greene, & House, 1977). This "egocentric" tendency may partially spring from the fact that personal preferences are complex yet readily accessible, and people adjust insufficiently from that initial belief when considering others (Epley, Keysar, Van Boven, & Gilovich, 2004). The ticket-payer can generate her preference easily (e.g., "it is so much easier to just mail a check than to go to the courthouse"). But possible corrections are either less likely to come to mind or unlikely to be sufficiently weighted, and so they fail to fully represent the perspective of the ticket-contester (e.g., "it is aversive to pay a fee that is undeserved"). Because it is difficult to perceive the forces that influence the preferences of others, people emphasize the easy-to-perceive forces that influence their own preferences.

These tendencies are potentially open to additional error as internal evaluations become more complicated, and therefore harder to estimate in others. The most basic form of such complexity might be in the simultaneous experience of both positive and negative emotions (Larsen, McGraw, & Cacioppo, 2001). Graduating seniors, for example, often have a somewhat mixed view of their graduation. On the one hand, there is the happiness of achieving a life goal, but on the other hand, there is the sadness of parting ways with friends. Such an experience can therefore be anticipated to have a bivariate emotional profile. College graduations may be extreme in this way that is, in producing mixed emotions. But many experiences may be characterized as at least partially mixed, even though they may tend to be perceived as primarily varying on a single dimension.

Consider the person anticipating the Rio De Janeiro vacation. They are likely aware of their own mixed emotional forecast, recognizing that even if there are many positives to the trip, there will still be some negatives as well (e.g., the effort to get there). Nevertheless, when evaluating the preferences of someone else, those negative attributes might not feel as intuitively compelling as the positive attributes. Any product or experience that is sought out, sold, or offered will presumably be identified as being generally positive. Therefore, when predicting the vacation experience of another person, the initial evaluation will not be purely egocentric (focusing on the personally mixed feelings), but rather, focused on the core representation of the experience itself (i.e., "vacations are enjoyable"). Consistent with this general possibility, recent research suggests that people maintain a simplified representation of other people's preferences. For instance, if someone is said to enjoy city vacations, people assume that same person would not enjoy a superficially dissimilar beach vacation (Barasz, Kim, & John, 2016).

People assess the representativeness of an experience by comparing it with an "ideal type" or a reference point of that experience (Rosch, 1975). The core representation reflects the seemingly

most probable, or most intuitively compelling, representation within the category of the experience (Kahneman & Tversky, 1972, 1982; Rosch, 1975). A sparrow is an intuitively compelling representation of birds, Linda is a compelling representation of feminists, and enjoyment is a compelling representation of a trip to Rio De Janeiro. Representativeness serves as a heuristic to allow people to answer a difficult question with a related (but not identical) easier one (Kahneman & Frederick, 2002; Kahneman & Tversky, 1972; Tversky & Kahneman, 1981). Relying on the intuitively compelling representation makes judgments easier, but it also makes them prone to error. Some birds are penguins, some feminists are bank tellers, and some trips to Rio De Janeiro are not very enjoyable. Judgments based on core representations necessarily fail to consider the broader distribution of experiences, so such judgments are likely to be more one-dimensional, uniformly valenced, and, therefore, biased. The Rio traveler may be aware of her own misgivings, but she may judge the experience of someone else as adhering closely to that prototypical representation of a visit to Rio.

People start with intuitively compelling representations, but they may still try to adjust their estimates to capture the integration of offsetting information. That process has been described in the operation of intuitive confidence (Simmons & Nelson, 2006). That theory holds that people make judgments by generating an initial judgment, and then, based on how easily that intuition comes to mind, decide how much it needs to be corrected. The model is an extension of other efforts to characterize how people trade off between fast and compelling initial thoughts and the more detailed and complicated reasoning that follows (e.g., System 1 vs. System 2; Kahneman & Frederick, 2002). For example, sports gamblers are often tasked with judging a prospect with strong intuitive value, but also strong need for correction. A single game may feature an obviously stronger team (say, the Warriors playing the Knicks in the NBA), but then also a point spread designed to equate the two options. Gamblers evaluate the prospect by forming an initial intuition (i.e., that the Warriors will beat the Knicks), and because that intuition is held with great confidence, fail to fully account for the accompanying point spread. Accordingly, although point spreads should make it equally likely for a bet on either team to pay-off, most bets are placed on the intuitive winner (Simmons & Nelson, 2006).

Similarly, predictions about the evaluations of others nearly always start with an easy intuition. Vacations are supposed to be fun, restaurant meals are supposed to be tasty, and parties are supposed to be festive. Accordingly, the first intuition about any of those will come to mind quickly, and the intuition will be positive. Furthermore, because many negative features are not an automatic part of evaluating a positive experience, they are less likely to be applied to correct those initial intuitions (Simmons & Nelson, 2018).

This line of prediction goes beyond the basic egocentric projection findings. Whereas existing research emphasizes that people overestimate how much their personal preferences are shared, we predict that they will think that other people have more intense evaluations of the same stimuli. The Rio De Janeiro traveler will not only overestimate the number of people who will choose Rio De Janeiro over Rome as a destination, but they will also believe that other people will enjoy Rio De Janeiro more than they will themselves. In this way, our predictions also follow in the tradition

of research on pluralistic ignorance (Katz & Allport, 1931; Miller & McFarland, 1987). In a prototypical investigation of pluralistic ignorance, the behavior of self and other are the same (e.g., both people choose not to ask a question at the end of a difficult lecture), but people draw different inferences about internal mental states (e.g., students in a classroom personally stay quiet because they are embarrassed to ask something foolish, but they think other students stay quiet because they understand the lecture). Thoughts about the valuations of others may follow a similar pattern. Although people have insight into their own mixed feelings (e.g., they choose to visit Rio De Janeiro despite not liking beaches), they cannot see that in others, and therefore assume that those feelings might be small or absent in others.

Although our prediction that people perceive others to experience things more intensely is new, the prediction is consistent with some related findings. Perhaps most notable about those findings is that they describe seemingly related behaviors, but they are hypothesized to derive from largely distinct mechanisms. For example, people believe that on average, other people will be willing to pay more than they will themselves (Frederick, 2012). When asked how much they would be willing to pay for 8 oz. of smoked salmon, people said that they would personally spend about \$3.00, but that they thought others would spend about \$6.67. That tendency extends across traditional products like smoked salmon, but also valuations for idiosyncratic offerings like a pill that makes you 2 in. taller (\$191 vs. \$895). This finding is suggested to arise in part from beliefs about how other people value money.<sup>1,2</sup> Similarly, research on the endowment effect has suggested that potential buyers overestimate how much other buyers would be willing to pay for identical goods (Kurt & Inman, 2013). This finding is suggested to arise from empathy gaps: Even when in the role of a buyer, people fail to realize how difficult it is for other buyers to give up their money to obtain a good. People also seem to think that others are getting more enjoyment out of social life, by attending more parties and events (Deri, Davidai, & Gilovich, 2017; Kahneman & Tversky, 1972; Tversky & Kahneman, 1973). This finding, yet again, is thought to arise from a different mechanism, in this case that people think of atypical exemplars rather than a true average. Finally, people also believe that, given the same product (e.g., a subscription to *Popular Science*), other people will use it more often (Ziano & Villanova, 2019). That finding is hypothesized to derive from people chronically overestimating the dispositional materialism of others. Though all of these mechanisms (valuation of money, empathy gaps, salient standards of comparison, or deeply different underlying psychologies) certainly play a role in comparative judgments of valuation, we suggest that they can all potentially be explained by the same underlying mechanism: People simply believe that other people have more uniformly valenced experiences.

Just as it is worth identifying some related findings that seem to be explained by our theorizing, it is just as important to point out a domain that would seem to show an opposing pattern of results. Research on empathy gaps frequently finds that people *underestimate* the experiences of others relative to themselves. People underestimate the experiences of others who are in a different physiological, emotional, or social situation (Nordgren, Banas, & MacDonald, 2011; Van Boven, Dunning, & Loewenstein, 2000; Van Boven & Loewenstein, 2003). For example, when people are not hungry or thirsty themselves, they tend to underestimate the

hunger and thirst of lost hikers in the woods (Van Boven & Loewenstein, 2003).

We briefly detail some salient differences between the different contexts, and then return to this topic in the General Discussion. The bias we document in this research is distinct from the egocentric empathy gaps or projection bias in several ways. First, research on empathy gaps primarily focuses on shortcomings in emotional perspective taking when self and other are in different psychological states. In fact, that bias is frequently contrasted with smaller or absent self-other differences when people are in similar psychological states. In our research we focus on people making judgments over the same stimuli about other participants who are drawn from the same population. The second distinction is one of research focus. Whereas research on empathy gaps focuses largely on egocentric self-projection, our research is largely focused on common shared representations of the stimuli. There is considerable evidence for projection in our studies, but we focus on effects that persist over and above that tendency. Third, and perhaps most interestingly, some of the discrepancies between our research and empathy gap research would appear to be driven by idiosyncrasies in the measurement of preferences. We return to the last point in the General Discussion.

Our article aims to document overestimation of others' preferences and valuations, establish its robustness across domains, and identify some evidence for our explanation. We first report evidence that people overestimate the preferences and valuations of others, both in a controlled environment (Study 1) and in the field (Study 2), and across products and experiences varying in expense,

<sup>1</sup> To be sure, Frederick (2012) makes clear that there is unlikely to be one singular explanation, but actually does claim to rule out the possibility that people overestimate the enjoyment of others more generally. That claim is based on evidence from one experiment (Study 3) showing that people did not believe that other people would enjoy experiences more than they would themselves. We conducted a direct replication of the original study with a much larger sample ( $N = 1,943$ , see Study S1 in the online supplemental studies and materials). We found that for a variety of products people believe others are both willing to pay more (80.5%,  $p < .001$ ) and enjoy the product more (57.5%,  $p < .001$ ) than they would themselves. The original would appear to be a false negative. Moreover, using a continuous scale as in our studies rather than the binary choice in Frederick's (2012) Study 3, we again find that people consistently believe that others are both willing to pay more and enjoy the product more than they would themselves. It may still be the case that the overestimated valuation is larger for willingness-to-pay (WTP) than for enjoyment—something that we frequently, but not uniformly find in our experiments—a possibility consistent with the idea that being asked about WTP is a very strong indicator that an evaluation should be primarily positive. Details of our investigation can be found in the online supplemental materials.

<sup>2</sup> Another related theory holds that, conditional on having the same amount of money, people perceive others as having less buying power than they do themselves (Polman, Effron, & Thomas, 2017). This intriguing possibility offers an imperfect account for the present findings in two respects. First, the explanation could partially account for overestimation of WTP (because people believe that others have to spend more to acquire the same outcome), but would not account for overestimation of enjoyment (because people do not believe that other people have a limited ability to dole out enjoyment points). Second, there is some reason to be cautious in interpreting the evidence underlying the Polman, Effron, and Thomas (2017) finding. We conducted a direct replication of the experiment underlying their primary claim, but observed equivocal results (see Study S2 in the online supplemental materials for details). We are reluctant to give a fuller accounting for the Polman et al. (2017) finding until the empirical foundation is easier to interpret.



practicality, and familiarity (Study 3). We show that the bias persists across measures of valuation that have different psychological properties (Study 4). This overestimation appears not only when self and other preferences are evaluated separately, but also when they are asked of the same person (i.e., within-subjects; Study 5). Furthermore, when evaluating negative stimuli, people overestimate the dislike of others relative to themselves, suggesting that people assume that others have more intense experiences rather than merely more positive experiences (Study 6). Moreover, this bias is not merely due to an artifact of how we measure preference (Study 7). Consistent with our theory about judgments based on core representations, the bias is reduced substantially when people make judgments about others' preferences for a uniformly positive stimulus (Studies 8A–8B). We then introduce a novel paradigm to investigate whether the effect operates through changes in the standard of comparison, by asking people to consider someone with an identical valuation on one measure, and then estimating their response on another. Consistent with the predictions of intuitive confidence, valuation overestimation persists even in this restricted condition, leading to a distinctive error in rationality (Studies 9A–9B). We propose that this overestimation bias arises because people's predictions about others' valuations of an experience reflect the core representations of the experience, whereas people's own valuations of the same experience involve consideration of broader dimensions of the experience and thus, are more moderate. Supporting this idea, when asked to first consider a broader distribution of others (both those who like and dislike the experiences), this overestimation is attenuated, and at times, eliminated (Study 10).

For all studies, we determined a stopping rule for data collection prior to data collection. For Study 2 (the field study), we collected responses from as many participants as possible in the month of data collection. For Study 8A, we aimed to collect responses from 800 participants. We preregistered all other studies before beginning data collection. We report all data exclusions, all manipulations, and all measures (Simmons, Nelson, & Simonsohn, 2013). Data, R code, and preregistrations for all studies reported in the main text and the additional supplemental materials are available at <https://osf.io/b8253/>. The institutional review boards of the authors' universities approved all of the following studies reported in this research.

### Analysis Plan

Some of our studies investigated self-other differences in evaluations between-subjects (Studies 1, 3, 4, 6, and 8A–8B), whereas other studies investigated self-other differences in evaluations within-subjects (i.e., participants reported both their own evaluation and their estimates of others' evaluations; Studies 2, 5, 7, and 9A–10). For between-subjects designs, we simply compared self and other evaluations using independent samples *t* tests. When examining willingness-to-pay (WTP), we winsorized self and other WTP at the 95th percentile prior to conducting analyses in order to reduce the influence of extreme responses.

For within-subjects designs, our analysis was a bit more complex. Because we were concerned with the *difference* between self and other evaluations, our analyses for these within-subjects studies were conducted as follows. We calculated the difference between other evaluations and self-evaluations (other—self) and

tested whether that difference was different from 0. We report one-sample *t* tests using these differences. Furthermore, when examining WTP, we winsorized the *difference* between self and other WTP at the 2.5th and 97.5th percentiles prior to conducting analyses in order to reduce the influence of extreme responses (those in which people reported a large difference between self and other evaluations). We report one-sample *t* tests using these winsorized self-other differences and report raw means and standard deviations for illustration.

### Study 1: Overestimating the Positive Experience of Others

Do people believe that others derive more enjoyment from positive experiences? Even the most positive event, we suggest, will frequently come with a slightly mixed experience. We theorize that this nuance is difficult to apply in judgments of others, so people will estimate others' experience as more uniformly positive than their own. We predict that this intuition guides their judgments about others' valuation and utility. As previously mentioned, Frederick (2012) has documented a supportive finding in showing that people overestimate how much other people are willing to pay for products and experiences. In Study 1, we test whether this tendency extends beyond estimation of WTP to assessment of others' liking. For either themselves or for someone else, participants reported either WTP or anticipated enjoyment. We predicted that people would think others would not only pay more for the movie, but also that others would expect to enjoy the movie more than they would.

### Method

**Participants and design.** Participants ( $N = 1,653$ ) recruited by Luth Research completed an online study for payment. Participants were randomly assigned to a 2 (person: self or other)  $\times$  2 (Dependent Variable [DV]: enjoyment or WTP) between-subjects design. Only participants who answered the attention check correctly ( $N = 1,311$ ) were included in the analyses (preregistration: <https://aspredicted.org/v8bz7.pdf>).

**Procedure.** Participants read a description of an upcoming movie (*Dog Days*), and watched the trailer for the movie. Then, participants were randomly assigned to answer questions about either themselves (*self* conditions) or another participant who was completing the survey (*other* conditions). Participants reported either WTP to watch the movie or anticipated enjoyment of the movie.

For WTP, participants indicated how much they (*self* condition) or another person taking the survey (*other* condition) would be willing to pay for a ticket to see the movie in dollars. For enjoyment, participants rated how much they (*self* condition) or another person taking the survey (*other* condition) would enjoy the movie on an 11-point scale (0 = *would not enjoy at all*; 10 = *would enjoy very much*).

### Results and Discussion

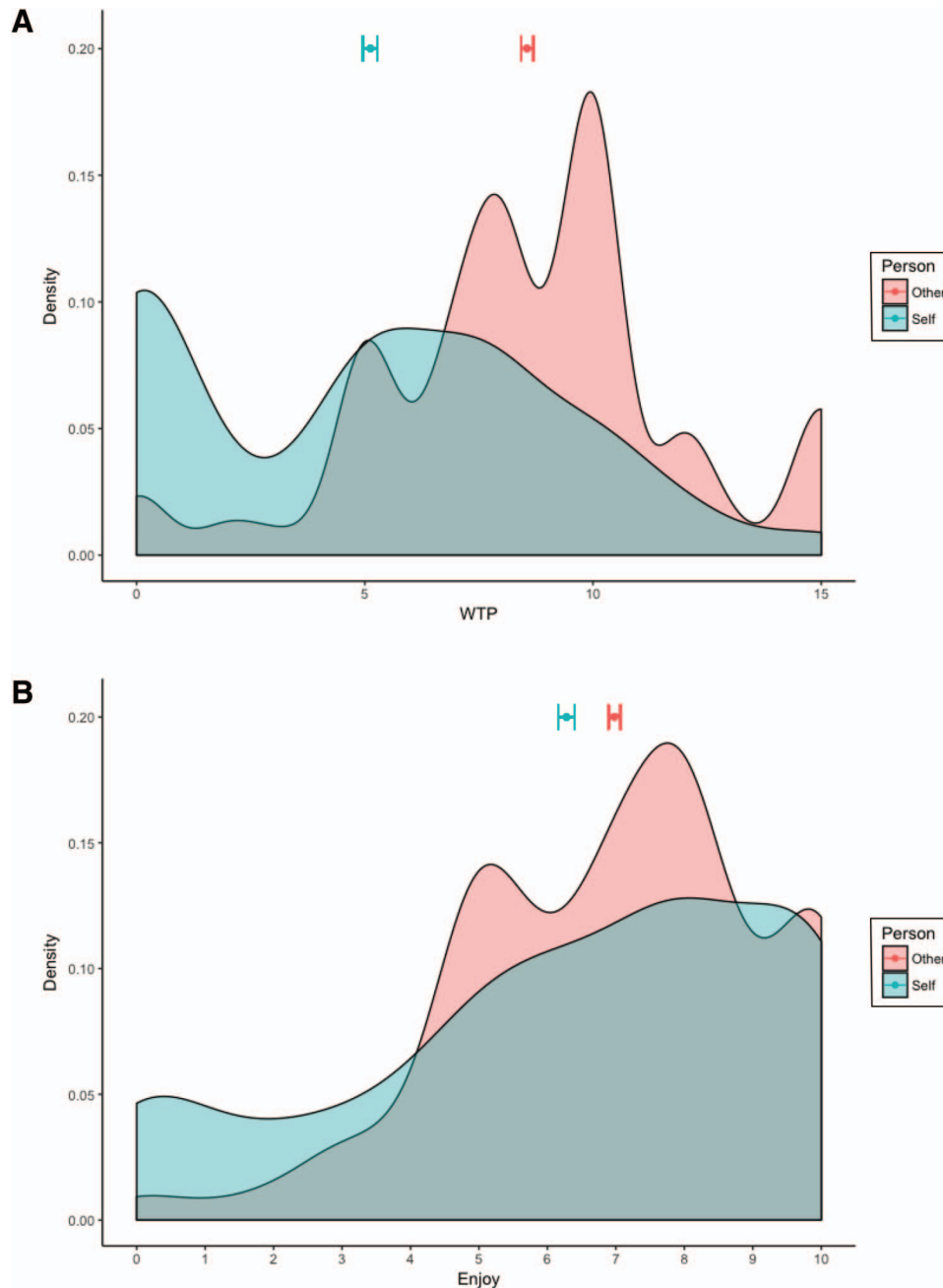
Replicating Frederick's (2012) finding, participants expected that others would be willing to pay significantly more for the movie ticket ( $M = \$8.56$ ,  $SD = \$3.39$ ) than they themselves would

pay ( $M = \$5.12$ ,  $SD = \$4.13$ ),  $t(676) = 11.92$ ,  $p < .001$ ,  $d = .92$  (see Figure 1A for a density plot of WTP).

Importantly for our investigation, people also overestimated others' enjoyment more generally, believing that others would enjoy the movie significantly more ( $M = 7.98$ ,  $SD = 2.21$ ) than they themselves would ( $M = 7.28$ ,  $SD = 3.09$ ),  $t(630) = 3.28$ ,  $p = .001$ ,  $d = .26$  (see Figure 1B for a density plot of enjoyment).

## Study 2: Field Validation

Study 1 provided initial evidence that people generally tend to overestimate others' preferences for a movie. One limitation of Study 1 is that participants in this study evaluated a movie that they had not chosen to watch. In contrast, when estimating how much another person might enjoy the movie, participants may have



**Figure 1.** Density plots for self and other. Means and error bars for self and other are displayed above the density plots. Panel A shows plots for self and other WTP. Panel B shows plots for self and other enjoyment. We note that other is an estimate of the average and thus, we should not expect the two distributions to look the same (e.g., we should not expect that many people would estimate others, on average, to like the movie at a 0, but there may be several people who report liking the movie at a 0 themselves). Nevertheless, we display the density plots for a richer illustration of the data. See the online article for the color version of this figure.

called to mind a person who wanted to watch the movie. Although we chose to ask participants to consider “another participant [. . .] completing this survey” to prevent this concern, it is still possible that participants visualized another person who would choose to watch the movie. The participants themselves, however, may not have been interested in watching the movie in question. This discrepancy could have artificially led to the overestimation of others’ liking for the movie. To address this possibility, in Study 2, we conducted a field study at a local movie theater to investigate whether overestimation of others’ WTP and enjoyment would persist in a setting in which people and the referent others both chose to watch a given movie.

## Method

**Participants and design.** Groups of customers ( $N = 274$ ) were approached to complete a short survey. Each group of customers were randomly assigned to a 2 (person order: self first or other first)  $\times$  2 (enjoyment: self or other)  $\times$  2 (DV order: WTP first or enjoyment first)  $\times$  2 (WTP: self or other) mixed-factors design with WTP as the only within-subjects factor.

**Procedure.** We collected data at a local movie theater in Oakland, California from 6 p.m. to 9:30 p.m. on four Wednesdays in Spring 2016. Every Wednesday the theater hosts a “Karma Cinema Night” during which 20% of all payments are donated to a local charity, and customers can choose to pay any price that they want, including \$0. The theater featured four movies on each Karma Cinema Night. The pay-what-you-want payment structure allowed us to obtain data on customers’ voluntary payments in an incentive-compatible setting.<sup>3</sup>

Four research assistants assumed roles as theater staff, greeted each group of customers as they entered the theater, and then asked them to complete a short questionnaire on an index card. As research assistants provided the card, they verbally instructed customers “Today you can pay what you want. Twenty percent of your payment goes to [a local charity name and the description of the charity].” This instruction was also provided on the questionnaire.

Each questionnaire consisted of five questions: (a) which movie they would like to watch, (b) the size of their group, (c) how much they plan to pay per person for a movie ticket, (d) how much they thought another customer who would watch the movie today would pay per person for a movie ticket, and (e) how much they thought they or another customer who would watch the movie today would enjoy the movie (randomly assigned to either report self or predict others’ enjoyment). We randomized whether participants reported enjoyment before or after a payment question, and whether participants reported their own payment first or their estimates of another customer’s payment. Thus, there were four versions of the questionnaire: (a) other payment, other enjoyment, self payment; (b) other enjoyment, other payment, self payment; (c) self payment, self enjoyment, other payment; and (d) self enjoyment, self payment, other payment (see Figure S1 in the online supplemental materials for the sample questionnaires).

Only one person from each group completed the questionnaire. After the person completed the questionnaire, they submitted it to the cashier along with their payment for the movie tickets. The cashier then recorded an invoice number on the back of each questionnaire submitted. Therefore, in addition to reports of self

WTP, we also tracked the actual sales records of each group of customers. The sales receipt recorded information about how much was spent on tickets as well as how much was spent on food and beverages.

## Results and Discussion

Seventeen groups of customers failed to complete enough data for any analysis. That is, they did not complete the enjoyment question nor did they complete payment information that would allow us to compare self-payment versus other-payment. Thus, our dataset contains the data from the remaining 257 groups of customers. For each analysis, our sample size depended on which subset of the 257 groups completed the measure(s) necessary.

We first sought to replicate the overestimation of others’ WTP in this field setting. We were able to analyze WTP in two different ways: (a) based on the customer invoice, and (b) based on the customer’s written response. As we describe next, our results were robust to either specification of WTP. Our first WTP analysis examined how much customers actually paid per person for the movie based on their invoice and compared it to how much they estimated that others would pay per person for the movie. Customers whose invoices were unrecorded were excluded from this analysis. Two-hundred and two groups of customers were included in this analysis. Consistent with previous research, people thought that other customers ( $M = \$7.01$ ,  $SD = \$2.79$ ) would pay more than they themselves did ( $M = \$6.03$ ,  $SD = \$3.45$ ), one-sample  $t(201) = 4.75$ ,  $p < .001$ ,  $d = .33$ . There were no interactions with either person order or DV order for these effects.

Our second WTP analysis compared customers self-reported payments with their perceptions of others’ payments. Two hundred thirty-six groups of customers were included in this analysis. Some customers had incorrectly reported how much they paid per ticket. A potentially common mistake would be a customer reporting the total price paid (e.g., \$10 for two tickets) rather than the per ticket payment (e.g., \$5 per ticket). Because that error is more likely to apply to self-payments than other-payments, the comparison using reported payments is somewhat conservative. Nevertheless, we found that the results from self-reported payments were consistent with the results from customers’ invoices: People thought that other customers would pay more ( $M = \$7.11$ ,  $SD = \$3.10$ ) than they reported paying themselves ( $M = \$6.21$ ,  $SD = \$3.57$ ),

<sup>3</sup> Though Study 2 was intended to test whether people overestimated others’ WTP and enjoyment in an incentive-compatible setting, voluntary payments, such as this pay-what-you-want payment structure, may reflect more than just people’s WTP. For instance, if payments in this setting also reflect customers’ concerns for appearing kind and generous, such a concern might drive customers to *underestimate* others’ payments compared with their own. In fact, research has shown that people believe that they will be more generous than others in charitable contexts (Epley & Dunning, 2000). On the other hand, more recent research examining voluntary payments—in pay-what-you-want and pay-it-forward settings—provides evidence that people believe that others will pay more than they themselves do even in these settings (Jung, Nelson, Gneezy, & Gneezy, 2014). Thus, Study 2 aimed to clarify whether people overestimate others’ payments in a voluntary payment setting. Furthermore, we note that our primary investigation in this study (self-other differences in people’s estimates of anticipated enjoyment) should be uninfluenced by this payment structure.

one-sample  $t(235) = 4.45, p < .001, d = .29$ . There were no interactions with either person order or DV order for these effects.

Next, we considered how people estimated their own anticipated enjoyment of the movie relative to the anticipated enjoyment of other customers. Two hundred forty-one groups of customers completed responses for this measure and were included in this analysis. replicating Study 1, people believed that other customers ( $M = 6.44, SD = .95$ ) would enjoy a movie more than they themselves would ( $M = 6.12, SD = 1.14$ ),  $t(239) = 2.37, p = .019, d = .31$ . There were no interactions with person order for these effects. There was, however, an unpredicted interaction between enjoyment and DV order,  $F(1, 237) = 4.16, p = .042$ . Though an order effect was not predicted, it is worth noting that: (a) this particular order effect is not driven by those who overestimate others' WTP and subsequently overestimate others' enjoyment of the same movie, and (b) this interaction is not robust. In later studies, we find that people believe that others enjoy products more than they themselves would even after reporting WTP.

### Study 3: Testing Overestimation Bias With Different Types of Products

Studies 1 and 2 provide converging evidence in lab and field settings that people exhibit an overestimation bias such that they believe that other people have more positive evaluations of movies than they themselves do. That bias persists whether measured with a value-based measure (WTP) or a subjective enjoyment scale. Though our claim is intended to be more general, our stimuli were necessarily narrow. In Study 3, we aimed to broaden our investigation to a number of different products. Although it is impossible to randomly sample all existing products and experiences, we selected exemplar products that varied on three dimensions that might be consequential—specifically, whether the product was hedonic or utilitarian, expensive or cheap, and whether its market price was easy or difficult to predict.

### Method

**Participants and design.** Participants ( $N = 802$ ) recruited from Amazon's Mechanical Turk (MTurk) completed an online study for payment. Participants were randomly assigned to a 2 (person: self or other)  $\times$  2 (DV: WTP or enjoyment)  $\times$  8 (product) mixed-factors design with product as the only within-subjects factor. Only participants who answered the attention check correctly ( $N = 771$ ) were included in the analyses (preregistration: <https://aspredicted.org/wx6m9.pdf>).

**Procedure.** As in Study 1, participants were randomly assigned to answer questions about either themselves (*self* conditions) or other participants taking the same survey (*other* conditions). Participants reported either WTP or enjoyment for eight different products.

Participants in the *self* conditions were asked to: (a) indicate their WTP for each product, or (b) rate how much they would enjoy having each product on a 101-point scale (0 = *not at all*; 100 = *extremely*).

Participants in the *other* conditions were first asked to consider all of the other participants taking the survey on MTurk. Participants were then asked to: (a) indicate their best estimate of those participants' average WTP for each product, or (b) rate how much

they thought those participants, on average, would enjoy having each product on a 101-point scale (0 = *not at all* to 100 = *extremely*). The products were: (a) a treadmill, (b) a 2018 Toyota Prius, (c) a movie ticket, (d) a round-trip flight to Paris, (e) a stick of butter, (f) a bottle of rose water, (g) a personal driver and limousine for a month, and (h) a tube of toothpaste.

### Results and Discussion

Replicating Frederick (2012), across the eight products, participants estimated that others would be willing to pay more for the products than they themselves would, aggregate  $t(374) = 3.26, p = .001, d = .34$  (see Table 1 for results by product).<sup>4</sup>

Importantly, across the eight products, participants also generally believed that others would enjoy the products more than they themselves would, aggregate  $t(393) = 4.70, p < .001, d = .47$  (see Table 2 for results by product). These results suggest that overestimation of others' valuation and utility is a broader phenomenon across different categories of products. Although the size of the effect varies across products, throughout this paper, we replicate the robust general tendency that people believe others will pay more for products and enjoy products more than they themselves would.<sup>5</sup>

### Study 4: Willingness-to-Wait

People have a general intuition that other people like things more than they do themselves. Is this bias specific to WTP and enjoyment, or does it manifest as a general intuition of others' preferences? In Study 4, we examine whether this overestimation bias emerges with yet another commonly used valuation metric that has different psychological properties from WTP and enjoyment: willingness-to-wait.

### Method

**Participants and design.** Participants ( $N = 1,228$ ) recruited from MTurk completed an online study for payment. Participants were randomly assigned to a 2 (person: self or other)  $\times$  4 (prod-

<sup>4</sup> Aggregate scores were determined by computing Z-scores for each product and averaging all Z-scores. An independent samples *t*-test was then conducted to compare these average Z-scores in the self versus other condition. For WTP, responses were winsorized before Z-scoring.

<sup>5</sup> As can be seen in Tables 1 and 2, the overestimation biases for WTP and enjoyment do not always appear for the same products. For example, participants overestimated others' liking for a bottle of rose water, but, while directionally consistent, their estimation of others' WTP did not reach significance. The direction even reverses for treadmills. There are a few possibilities for how to think about this. The easiest is that measures of valuation are imperfect, so it is entirely reasonable that the same stimuli could be evaluated differently with different measures. Another possibility is that there are specific products (e.g., rose water and treadmills) that operate distinctively on the more general phenomenon under investigation. This might be potentially fruitful in future research, but we are reluctant to overinterpret based on the limited number of observations available. Finally, a third approach might be to look to other existing findings comparing responses on WTP and subjective scale measures (e.g., Moon & Nelson, in press) for theoretically driven predictions about how and when core representations might influence different evaluations. That might also have potential, but as of this writing, we could not identify a sufficiently unifying explanation that we wanted to stand by. Future researchers might be more successful.



Table 1  
WTP for Self Versus Other in Study 3

| Product                   | Self ( <i>n</i> = 187) | Other ( <i>n</i> = 189) | <i>t</i> | Cohen's <i>d</i> |
|---------------------------|------------------------|-------------------------|----------|------------------|
| Treadmill                 | \$279.34 (\$222.90)    | \$373.98 (\$255.19)     | 3.83***  | .40              |
| Toyota Prius              | \$19796.24 (\$9035.65) | \$22506.54 (\$7668.10)  | 3.02**   | .31              |
| Movie ticket              | \$10.22 (\$3.72)       | \$11.17 (\$3.66)        | 2.50*    | .26              |
| Flight to Paris           | \$901.12 (\$520.98)    | \$981.41 (\$503.74)     | 1.52     | .16              |
| Stick of butter           | \$1.71 (\$1.23)        | \$1.84 (\$1.29)         | 1.03     | .11              |
| Bottle of rose water      | \$6.75 (\$7.14)        | \$7.34 (\$6.64)         | .83      | .09              |
| Personal limousine driver | \$759.41 (\$909.86)    | \$782.95 (\$843.83)     | .26      | .03              |
| Tube of toothpaste        | \$2.95 (\$1.37)        | \$2.93 (\$1.24)         | -.10     | -.01             |
| Total                     |                        |                         | 3.26**   | .34              |

Note. Standard deviations are in parentheses. (When the effect is in the opposite direction, we indicate Cohen's *d* as negative to more easily illustrate that the effect is counter to our predictions.)

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

ucts) between-subjects design. Only participants who answered the attention check correctly ( $N = 953$ ) were included in the analyses (preregistration: <https://aspredicted.org/jw7f7.pdf>).

**Procedure.** Participants were randomly assigned to answer questions about either their own or another participant's willingness-to-wait for a product. Participants in the *self* conditions indicated how long they would be willing to wait to receive the product. Participants in the *other* conditions predicted on average, how many minutes other participants taking this survey would be willing to wait to receive the product.

The products were: (a) an early screening of the newest *Star Wars* movie, (b) a newly released virtual reality headset (Oculus Go), (c) an autographed jersey of Michael Jordan, or (d) a tour of the Burj Khalifa skyscraper in Dubai.

## Results and Discussion

Participants predicted that another participant in the same study would be willing to wait longer for the products than they themselves would, aggregate  $t(951) = 8.14$ ,  $p < .001$ ,  $d = .53$  (see Table 3 for results by product and Figure 2 for a density plot of the early movie experience).<sup>6</sup> These results suggest that the overesti-

mation bias appears more broadly in the estimation of others' values and preferences.

## Study 5: Within-Subjects

Studies 1–4 used between-subjects designs in which participants indicated either their own or others' valuation of products. It is possible that people do not actually overestimate others' valuation when they more directly compare it to their own. In Study 5, we test whether the observed overestimation bias still persists in a within-subjects design in which participants indicate their own WTP or liking as well as their estimation of others' WTP or liking.

## Method

**Participants and design.** Participants ( $N = 1,629$ ) recruited by Luth Research completed an online study for payment. Participants were randomly assigned to a 2 (person: self or other)  $\times$  2 (DV: WTP or enjoyment) mixed-factors design with person as the only within-subjects factor. Only participants who answered the attention check correctly ( $N = 1,439$ ) were included in the analyses (preregistration: <https://aspredicted.org/np7qd.pdf>).

**Procedure.** Participants evaluated a 12-piece box of Godiva Chocolatier Signature Chocolate Truffles. Participants in the WTP condition were asked to indicate their WTP for the box of chocolates as well as their estimate of the average WTP for other participants taking the survey. Participants in the enjoyment condition were asked to indicate their own enjoyment for the box of chocolates on a 101-point scale (0 = *not at all* to 100 = *extremely*) as well as their estimate of the average enjoyment of the other participants taking the survey. We counterbalanced the order of self-evaluations and other-evaluations.

## Results and Discussion

Even within-subjects, participants expected that others would be willing to pay significantly more for the box of

Table 2  
Enjoyment for Self Versus Other in Study 3

| Product                   | Self ( <i>n</i> = 193) | Other ( <i>n</i> = 202) | <i>t</i> | Cohen's <i>d</i> |
|---------------------------|------------------------|-------------------------|----------|------------------|
| Treadmill                 | 54.55 (30.89)          | 49.50 (22.54)           | -1.86    | -.19             |
| Toyota Prius              | 80.45 (25.19)          | 86.57 (14.56)           | 2.97**   | .30              |
| Movie ticket              | 60.48 (30.04)          | 74.84 (19.98)           | 5.62***  | .56              |
| Flight to Paris           | 81.96 (24.88)          | 85.09 (15.31)           | 1.52     | .15              |
| Stick of butter           | 25.05 (27.14)          | 32.04 (25.93)           | 2.62**   | .26              |
| Bottle of rose water      | 32.58 (29.50)          | 41.45 (22.86)           | 3.35***  | .34              |
| Personal limousine driver | 68.67 (32.31)          | 78.80 (20.32)           | 3.75***  | .37              |
| Tube of toothpaste        | 29.61 (28.41)          | 37.10 (28.60)           | 2.61**   | .26              |
| Total                     |                        |                         | 4.70***  | .47              |

Note. Standard deviations are in parentheses. (When the effect is in the opposite direction, we indicate Cohen's *d* as negative to more easily illustrate that the effect is counter to our predictions.)

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

<sup>6</sup> Similar to Study 3, aggregate scores were determined by Z-scoring the winsorized wait times by product. Because product was between subjects, the aggregate score reflected each Z-score in one variable. An independent samples *t*-test was then conducted to compare the Z-scores for self versus other.



Table 3  
Willingness-to-Wait in Minutes for Products and Experiences in Study 4

| Products   | Self                             | Other                             | <i>t</i> | Cohen's <i>d</i> |
|--|----------------------------------|-----------------------------------|----------|------------------|
| Early screening of the newest <i>Star Wars</i> movie | 42.40 (40.32)<br><i>n</i> = 119  | 99.72 (87.53)<br><i>n</i> = 127   | 6.52***  | .83              |
| An autographed jersey of Michael Jordan              | 98.77 (110.21)<br><i>n</i> = 123 | 160.93 (136.84)<br><i>n</i> = 122 | 3.92***  | .50              |
| A newly released virtual reality headset             | 47.90 (70.63)<br><i>n</i> = 116  | 89.53 (94.91)<br><i>n</i> = 116   | 3.79***  | .50              |
| A tour of the Burj Khalifa skyscraper in Dubai       | 60.73 (39.71)<br><i>n</i> = 117  | 73.21 (47.95)<br><i>n</i> = 113   | 2.15*    | .28              |
| Total  |                                  |                                   | 8.14***  | .53              |

Note. Standard deviations are in parentheses.

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

chocolates than they themselves would ( $M_{\text{Difference Score}} = \$3.21$ ,  $SD_{\text{Difference Score}} = \$5.38$ ), one-sample  $t(740) = 16.26$ ,  $p < .001$ ,  $d = .60$  (see Figure 3A).

Moreover, within-subjects, participants also expected that others would enjoy the box of chocolates significantly more ( $M = 80.44$ ,  $SD = 20.36$ ) than they themselves would ( $M = 75.89$ ,  $SD = 29.08$ ), one-sample  $t(697) = 5.55$ ,  $p < .001$ ,  $d = .21$  (see Figure 3B). These results confirm that the overestimation bias appears regardless of whether it is measured in a between-subjects or a within-subjects design.

### Study 6: Negative Experiences

Studies 1–5 provide robust evidence that people overestimate others' valuation of products. According to our theorizing, people estimate others' experiences to be more uniformly consistent with the core representation of the experience. For instance, chocolates generally represent a positive product. Thus, even though people might themselves have a more mixed valuation of chocolates (e.g., chocolates are tasty but fattening), they tend to think that others have a more uniformly positive consumption experience. By the

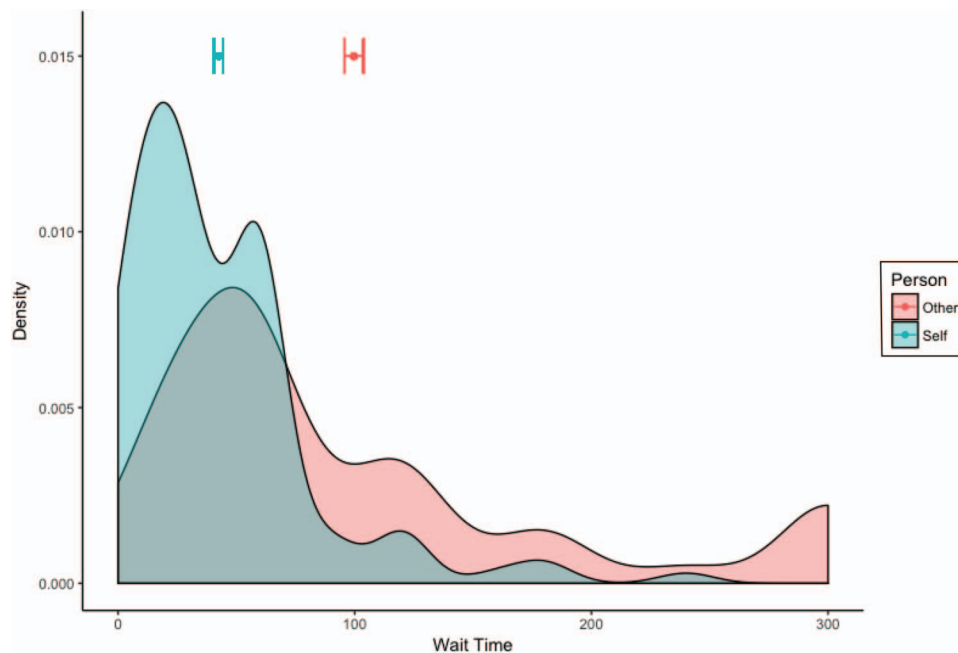
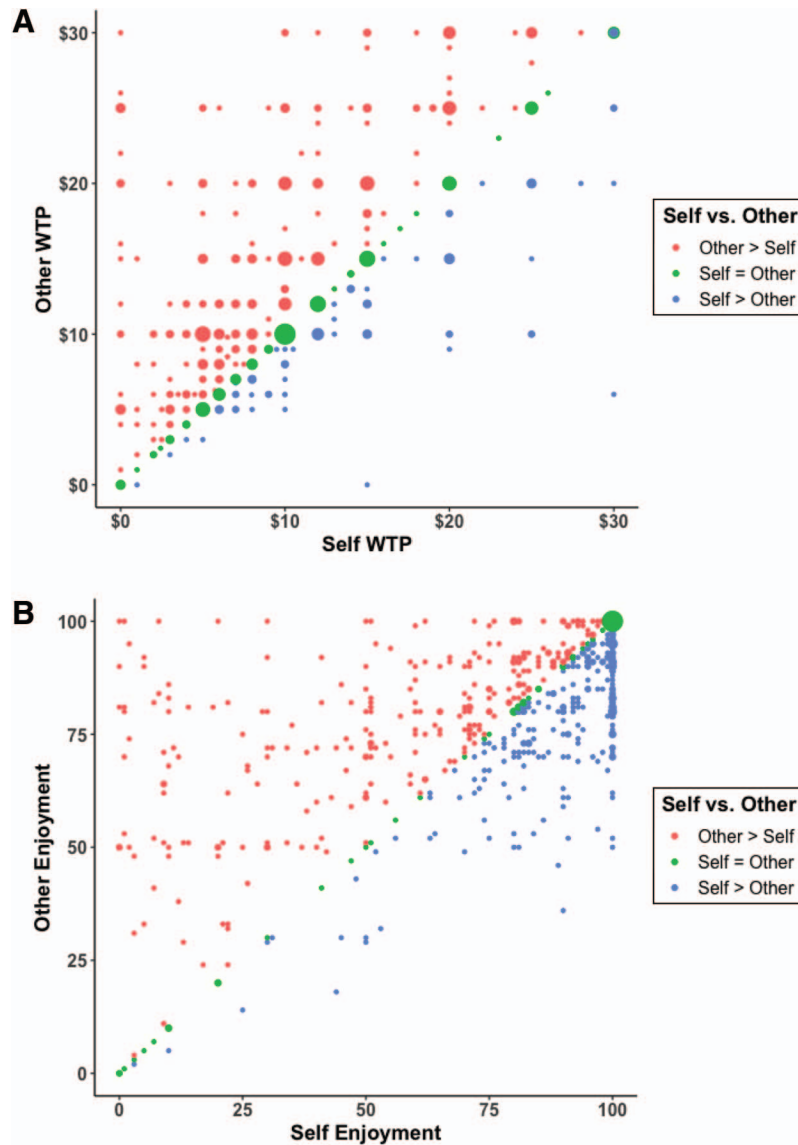


Figure 2. Density plots for self and other wait times for an early movie screening in Study 4. Means and error bars are shown above each graph. As with Figures 1A–B, we note that other is an estimate of the average and thus, we should not expect the two distributions to look the same (e.g., we should not expect that many people would estimate others, on average, to wait 300 min for the movie screening, but there were several people who reported wanting to wait 300 min themselves). Nevertheless, we display the density plots for a richer illustration of the data. See the online article for the color version of this figure.



*Figure 3.* Self and other preference for a box of chocolates in Study 5. Larger circles denote a greater number of observations at that data point. Panel A shows self and other WTP. For illustration, WTP for payments were winsorized at \$30 for both self and other (roughly 95th percentile). Panel B shows self and other anticipated enjoyment. WTP = willingness-to-pay. See the online article for the color version of this figure.

same logic, if the core representation of an experience is negative (e.g., washing dishes is annoying but productive), they should predict that others' experience will be more negative than their own. That is, if they believe that others tend to have more uniformly consistent consumption experiences, they should also think that others dislike negative experiences more than they would and that others would be willing to pay more to avoid negative experiences than they themselves would.

Alternatively, people might think that others are generally more positive about experiences. If so, then people should also think that others dislike negative experiences *less* than they would themselves. This is not an implausible alternative. For example, people think that they will need more compensation to

perform embarrassing acts than would others (Van Boven, Loewenstein, & Dunning, 2005). Because existing products tend to be considered positive (e.g., a trip to Paris) or neutral (e.g., a tube of toothpaste), Studies 1–5 cannot differentiate between these two explanations. Thus, in Study 6, we aim to disentangle these competing hypotheses by investigating whether people think that others would dislike negative experiences more or less than they themselves would.

## Method

**Participants and design.** Participants ( $N = 3,291$ ) recruited from MTurk completed an online study for payment. Participants

were randomly assigned to a 2 (person: self or other)  $\times$  2 (DV: WTP to avoid or dislike)  $\times$  8 (negative experiences) mixed-factors design with negative experiences as the only within-subjects factor. Only participants who answered the attention check correctly ( $N = 3,096$ ) were included in the analyses (preregistration: <https://aspredicted.org/7c8fv.pdf>).

**Procedure.** Participants were randomly assigned to answer questions about either themselves or about others. Participants in the *self* condition reported how much they would be willing to pay to avoid or how much they would dislike eight negative experiences on a 101-point scale (0 = *not at all* to 100 = *extremely*). Participants in the *other* condition predicted, on average, how much other participants would be willing to pay to avoid, or on average, how much other participants would dislike eight negative experiences on a 101-point scale (0 = *not at all* to 100 = *extremely*).

We used the eight negative experiences in Appendix J of Fredrick (2012). The negative experiences were: (a) eating four hard-boiled eggs over 15 min, (b) walking around your neighborhood with a t-shirt that says “Repent!,” (c) sharpening 100 pencils using a mechanical sharpener, (d) going without sleep tonight, (e) fasting for 4 consecutive days, (f) composing and reciting a poem about love for a paid task, (g) eating three large jalapeno peppers with no drink, and (h) shaving your head completely. The negative experiences were presented in a random order.

## Results and Discussion

Consistent with our suggested explanation, across the eight negative experiences, participants believed both that others would be willing to pay more to avoid the negative experiences than they themselves would, aggregate  $t(1,532) = 4.32, p < .001, d = .22$  (see Table 4 for results by experience), and that others would dislike the negative experiences more than they themselves would, aggregate  $t(1,560) = 4.71, p < .001, d = .24$  (see Table 5 for results by experience).<sup>7</sup>

These results support the notion that people believe that other people have more intense experiences, rather than simply more positive experiences. (Again, we do not mean “intensity” in the sense of amplified experiences [e.g., people do not believe that others have a more intense sensation of boredom when sharpening pencils with a mechanical sharpener] but instead in the sense that judgments of others’ experiences are not tempered by nonrepresentative attributes [e.g., people do not consider that others might find sharpening pencils could be relaxing].) This is consistent with our intuitive confidence theorizing of how people are influenced by the core representation of products and experiences.

Notably, this observed overestimation bias of negative experiences is in seeming contrast to well-documented egocentric empathy gaps and social projection (e.g., Van Boven et al., 2000). We think the difference is intriguing. As we allude to in the Introduction, we think there are a few possible explanations, but we give the most detail to the possibility of a measurement issue in the General Discussion.

### Study 7: Measuring Preferences

In our studies thus far, we consistently found that people overestimated others’ liking for positive experiences and others’ disliking for negative experiences. We interpret those findings as

suggesting that people show a general tendency to think that other people experience things more intensely. There is, however, some concern about a subtle measurement artifact. For each of the stimuli we used in the previous studies, there is a predominant valence and for each of the measures we collected, there is a unipolar scale that matches that valence. Consider, for example, the box of chocolates. Because chocolates are sold, even a chocolate-averse participant will still be aware that many other people like chocolates, so any evaluation will presumably be positive. Our measures mirror that. The subjective rating scale is anchored on the low end at *do not like at all* rather than *dislike very much*. Similarly, the lowest coherent WTP response is \$0 rather than a negative value, so that measure intrinsically has a unipolar scaling. It may be the case, therefore, that the effects we document are a product of the match between stimuli and scales, rather than a more general tendency in human judgment.

In Study 7, we try to address this potential confound by testing mismatched and bipolar rating scales. Perhaps, by allowing people to report mixed feelings toward a stimulus, they will stop seeing others as having more intense experiences.

## Method

**Participants and design.** Participants ( $N = 1,215$ ) recruited from MTurk completed an online study for payment. Participants were randomly assigned to a 2 (person: self or other)  $\times$  2 (experience: positive or negative)  $\times$  3 (scale: like, dislike, or bipolar) mixed-factors design with person as the only within-subjects factor. Only participants who answered the attention check correctly ( $N = 1,095$ ) were included in the analyses (preregistration: <https://aspredicted.org/d5hi4.pdf>).

**Procedure.** Participants in the *positive experience* conditions evaluated a 30-min back massage. Participants in the *negative experience* conditions evaluated a task of sharpening 100 pencils using a mechanical sharpener.

In the *like* conditions, participants evaluated how much they would like the experience as well as how much they thought the average other participant completing the study would like the experience on a 101-point scale ranging from 0 (*not at all*) to 100 (*extremely*). In the *dislike* conditions, participants evaluated how much they would dislike the experience as well as how much they thought the average other participant in the study would dislike the experience on a 101-point scale ranging from 0 (*not at all*) to 100 (*extremely*). In the *bipolar* conditions, participants evaluated how much they would like or dislike the experience as well as how much they thought the average other participant in the study would like or dislike the experience on a 101-point scale from  $-50$  (*dislike extremely*) to  $+50$  (*like extremely*) with 0 (*neither like nor dislike*) as the midpoint.

## Results

**Positive experience.** Replicating our previous results, participants thought other participants would like a massage more ( $M = 81.95, SD = 16.33$ ) than they themselves would ( $M = 76.86$ ,

<sup>7</sup> Aggregate scores were calculated in the same way as in Study 3 (see Footnote 3).

Table 4

*WTP to Avoid for Self and Other in Study 6*

| Negative experiences   | Self ( <i>n</i> = 758) | Other ( <i>n</i> = 776) | <i>t</i> | Cohen's <i>d</i> |
|--|------------------------|-------------------------|----------|------------------|
| Sharpening 100 pencils using a mechanical sharpener                | \$7.18 (\$11.84)       | \$9.79 (\$12.71)        | 4.17***  | .21              |
| Fasting for 4 consecutive days                                     | \$96.94 (\$141.54)     | \$123.27 (\$142.39)     | 3.63***  | .19              |
| Going without sleep tonight  | \$33.53 (\$48.21)      | \$41.53 (\$50.78)       | 3.17**   | .16              |
| Walking around your neighborhood with a t-shirt that says "Repent" | \$18.02 (\$30.18)      | \$22.47 (\$30.23)       | 2.89**   | .15              |
| Composing and reciting a poem about love                           | \$7.35 (\$13.59)       | \$8.98 (\$13.69)        | 2.33*    | .12              |
| Eating four hardboiled eggs over 15 min                            | \$6.61 (\$13.36)       | \$7.84 (\$12.33)        | 1.86     | .09              |
| Eating three large jalapeno peppers with no drink                  | \$34.61 (\$54.24)      | \$38.69 (\$49.76)       | 1.54     | .08              |
| Shaving your head completely                                       | \$187.08 (\$304.46)    | \$212.31 (\$298.07)     | 1.64     | .08              |
| Total  |                        |                         | 4.32***  | .22              |

Note. Standard deviations are in parentheses.

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

$SD = 28.73$ ), one-sample  $t(146) = 2.68$ ,  $p = .008$ ,  $d = .22$  (see Table 6 for a summary of results).

Importantly, this effect was consistent regardless of the scale used. Participants also thought that other participants would dislike a massage less ( $M = 14.16$ ,  $SD = 19.02$ ) than they themselves would ( $M = 23.52$ ,  $SD = 31.56$ ), one-sample  $t(190) = 4.97$ ,  $p < .001$ ,  $d = .36$ . And finally, using a bipolar scale, participants again thought that other participants would like a massage more ( $M = 35.37$ ,  $SD = 13.99$ ) than they themselves would ( $M = 27.12$ ,  $SD = 27.34$ ), one-sample  $t(192) = 4.56$ ,  $p < .001$ ,  $d = .33$ .

**Negative experience.** Replicating our previous results, participants thought other participants would dislike sharpening pencils more ( $M = 66.48$ ,  $SD = 26.06$ ) than they themselves would ( $M = 56.68$ ,  $SD = 30.15$ ), one-sample  $t(184) = 6.48$ ,  $p < .001$ ,  $d = .48$ .

As with the positive experience, this effect was consistent regardless of the scale used. Participants also thought that other participants would like sharpening pencils less ( $M = 23.53$ ,  $SD = 22.71$ ) than they themselves would ( $M = 28.41$ ,  $SD = 27.08$ ), one-sample  $t(178) = 4.10$ ,  $p < .001$ ,  $d = .31$ . And finally, using a bipolar scale, participants again thought that other participants would dislike sharpening pencils more ( $M = -21.09$ ,  $SD = 20.84$ ) than they themselves would ( $M = -12.63$ ,  $SD = 24.44$ ), one-sample  $t(199) = 6.34$ ,  $p < .001$ ,  $d = .45$ .

These results provide evidence that the overestimation bias appears regardless of how valuation and preferences are mea-

sured. This seems to rule out the potential artifactual explanation for the overestimation bias. Even when using mismatched or bipolar scales, people believe that others have more intense positive and negative experiences than they themselves do.

### Study 8A: Uniformly Positive Experience

Studies 1–7 provide evidence that people see others as having more intense experiences. We argue that this overestimation occurs because people tend to evaluate others' experiences as more uniformly consistent with the core representation of the experience, whereas the self's experiences are more mixed. If this is the case, then an experience that itself is uniform, rather than mixed, should attenuate (or possibly eliminate) this bias. That is, whereas the core representation of most stimuli masks the complexity of true experience, for a uniformly positive experience the core representation is largely accurate. Study 8A tests this prediction by comparing two prospects: (a) an experience of winning money (an experience that should be uniformly positive); and (b) a gamble with the possibility of winning money (an experience that is generally positive but may be mixed, i.e., by including the possibility of disappointment). By our core representation account, the uncertainty of the outcome might produce a mixed evaluation for one's self, but predictions of others' evaluations would be largely based on the upside of the prospect. Therefore, we predict that the

Table 5

*Dislike for Self and Other in Study 6*

| Negative experiences   | Self ( <i>n</i> = 799) | Other ( <i>n</i> = 763) | <i>t</i> | Cohen's <i>d</i> |
|--|------------------------|-------------------------|----------|------------------|
| Sharpening 100 pencils using a mechanical sharpener              | 47.53 (30.02)          | 55.11 (28.82)           | 5.09***  | .26              |
| Fasting for 4 consecutive days                                   | 78.99 (28.09)          | 85.37 (22.11)           | 4.97***  | .25              |
| Going without sleep tonight                                      | 78.61 (27.20)          | 77.50 (25.63)           | -.83     | -.04             |
| Walking around your neighborhood with a shirt that says "Repent" | 63.52 (33.68)          | 63.65 (26.20)           | .08      | .00              |
| Composing and reciting a poem about love                         | 52.14 (32.78)          | 53.96 (27.18)           | 1.20     | .06              |
| Eating four hardboiled eggs over 15 min                          | 44.33 (35.04)          | 49.23 (25.77)           | 3.14**   | .16              |
| Eating three large jalapeno peppers with no drink                | 73.28 (31.71)          | 76.72 (24.88)           | 2.38*    | .12              |
| Shaving your head completely                                     | 71.67 (34.02)          | 77.40 (22.37)           | 3.92***  | .20              |
| Total  |                        |                         | 4.71***  | .24              |

Note. Standard deviations are in parentheses. (When the effect is in the opposite direction, we indicate Cohen's *d* as negative to more easily illustrate that the effect is counter to our predictions.)

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



Table 6  
Self and Other Ratings of Preference Using Different Scales in Study 7

| Product                       | Like          |               | Dislike             |               | Bipolar             |                |
|-------------------------------|---------------|---------------|---------------------|---------------|---------------------|----------------|
|                               | Self          | Other         | Self                | Other         | Self                | Other          |
| Massage (positive)            | 76.86 (28.73) | 81.95 (16.33) | 23.52 (31.56)       | 14.16 (19.02) | 27.12 (27.34)       | 35.37 (13.99)  |
| Sharpening pencils (negative) | 28.41 (27.08) | 23.53 (22.71) | 56.68 (30.15)       | 66.48 (26.06) | -12.63 (24.44)      | -21.09 (20.84) |
|                               |               |               | One-sample <i>t</i> |               | One-sample <i>t</i> |                |
|                               |               |               | 2.68**              |               | -4.97***            |                |
|                               |               |               | -4.10***            |               | 6.48***             |                |

Note. Positive *t*-statistics indicate that other ratings are higher on the scale than self. Standard deviations are in parentheses.  
\*\*  $p < .01$ . \*\*\*  $p < .001$ .

overestimation bias will persist for the gamble but will be attenuated with a pure win.

### Method

**Participants and design.** Participants ( $N = 816$ ) recruited from MTurk completed an online study for payment. Participants were randomly assigned to a 2 (person: self or other)  $\times$  2 (prospect: gamble or pure win) between-subjects design. Only participants who answered the attention check correctly ( $N = 693$ ) were included in the analyses.

**Procedure.** Participants were randomly assigned to either evaluate a positive gamble or a positive experience of winning money. In particular, participants in the *gamble* condition were asked to imagine a gamble with a 50% chance of winning \$500 (and a 50% chance of winning nothing). Participants in the *pure win* condition were asked to imagine an experience of winning \$500. Participants were also randomly assigned to answer questions about either themselves or about others. Participants in the *self* condition reported how much they thought they would like the prospect on a 101-point scale (0 = *not at all* to 100 = *extremely*). Participants in the *other* condition were asked to consider another participant completing the survey. They predicted how much they thought this other participant would like the prospect on a 101-point scale (0 = *not at all* to 100 = *extremely*).

### Results and Discussion

We conducted a 2 (person: self or other)  $\times$  2 (prospect: gamble or pure win) between-subjects ANOVA on liking. There was a main effect of person such that people thought others would enjoy the prospects more than they themselves would,  $F(1, 689) = 10.10, p = .002$ . There was a main effect of prospect such that people thought the experience of winning \$500 would be more enjoyable than the 50% chance to win \$500,  $F(1, 689) = 136.19, p < .001$ . Most importantly, there was a significant Person  $\times$  Prospect interaction,  $F(1, 689) = 5.60, p = .018$ .

With the gamble, we replicated the effect from the previous studies: People thought that others would like the gamble more ( $M = 73.41, SD = 22.73$ ) than they themselves would ( $M = 67.00, SD = 23.96$ ),  $t(287) = 2.33, p = .020, d = .28$ . However, this effect was eliminated with the experience of winning \$500,  $t(402) = .15, p = .882$ . People thought that others would enjoy the experience of winning just as much ( $M = 95.22, SD = 11.08$ ) as they themselves would ( $M = 95.06, SD = 10.39$ ). This is compelling evidence for our core representation account: A uniformly positive experience does not display the same overestimation bias as a positive but mixed experience (a gamble).

### Study 8B: Replication With a Smaller Gain

One limitation with Study 8A is that we may have reached a ceiling effect with the pure win of \$500. Everyone anticipated enjoying the experience so much that there was no room left on the scale to differentiate self and other. That could have artificially led to the attenuated interaction when comparing the pure win with the gamble. Therefore, in Study 8B, we conducted a follow-up study to test our hypothesis with a more mildly positive experience—that is, with a smaller gain of \$10 instead of \$500.

## Method

**Participants and design.** Participants ( $N = 1,006$ ) recruited from MTurk completed an online study for payment. Participants were randomly assigned to a 2 (person: self or other)  $\times$  2 (prospect: gamble or pure win) between-subjects design. Only participants who answered the attention check correctly ( $N = 892$ ) were included in the analyses (preregistration: <https://aspredicted.org/6eh5n.pdf>).

**Procedure.** As in Study 8A, participants were randomly assigned to either evaluate a positive gamble or a positive experience of winning money. In particular, participants in the *gamble* condition were asked to imagine a gamble with a 50% chance of winning \$10 (and a 50% chance of winning nothing). Participants in the *pure win* condition were asked to imagine an experience of winning \$10. Participants were also randomly assigned to answer questions about either themselves or about others. Participants in the *self* condition reported how much they thought they would like the prospect on a 101-point scale (0 = *not at all* to 100 = *extremely*). Participants in the *other* condition were asked to consider another participant completing the survey. They predicted how much they thought this other participant would like the prospect on a 101-point scale (0 = *not at all* to 100 = *extremely*).

## Results and Discussion

We conducted a 2 (person: self or other)  $\times$  2 (prospect: gamble or pure win) between-subjects ANOVA on liking. There was a main effect of person such that people thought others would enjoy the prospects more than they themselves would,  $F(1, 885) = 48.49$ ,  $p < .001$ . There was a main effect of prospect such that people thought the experience of winning \$10 would be more enjoyable than the 50% chance to win \$10,  $F(1, 885) = 110.68$ ,  $p < .001$ . Most importantly, replicating Study 8A, there was a significant Person  $\times$  Prospect interaction,  $F(1, 885) = 13.66$ ,  $p < .001$ .

Decomposing the interaction, with the gamble, we replicated the effect from the previous studies: People thought that others would like the gamble more ( $M = 67.81$ ,  $SD = 19.26$ ) than they themselves would ( $M = 53.72$ ,  $SD = 25.74$ ),  $t(409) = 6.24$ ,  $p < .001$ ,  $d = .62$ . With the pure win, this effect was still significant (self:  $M = 84.62$ ,  $SD = 18.69$ ; other:  $M = 88.51$ ,  $SD = 17.69$ ),  $t(476) = 2.34$ ,  $p = .020$ ,  $d = .21$ , but as reflected in the significant interactions, the difference between self and other was attenuated for this prospect. This suggests that uniformly positive experiences exhibit attenuated overestimation as predicted by our core representation account.

### Studies 9A-9B: Others With the Identical Value or Preferences

Studies 1–8 provide evidence that people see others as having more intense experiences. We argue that this is largely driven by a tendency to evaluate others' experiences as more uniformly consistent with the core representation of the experience. Studies 9A–9B intended to add to this argument in two ways.

The first is a clarification of mechanism. The current findings are consistent with the idea that people perceive others as having experiences in line with the valence of those experiences' core

representations. The results are also consistent with an adjacent, but slightly different, account. That possibility suggests that people are simply recruiting a different (and biased) exemplar from the distribution, rather than misperceiving the distribution overall. When asked to consider others, it is possible that people initially think of a prototypical other rather than an average other. In concept, those could be the same, but when evaluating consumers of a box of chocolates, a prototypical other is perhaps more likely to be someone who buys chocolates regularly rather than to be someone who buys chocolates rarely. Because it is easier to observe others buying and liking than others not buying and not liking, those people will be more available in memory (Tversky & Kahneman, 1973). Subsequently, thinking of a prototypical consumer could lead to relatively more extreme judgments of other people's experiences, because a frequent buyer is likely to be more positive about chocolates. Recent research provides some evidence for this exceptional *prototypical other* account in comparative social predictions. When comparing their social life to those of others, people tend to think of someone with an exceptionally rich social life (Deri et al., 2017). Our own studies are consistent with a broadening of the same idea to goods and experiences. Just as people tend to think of other people with rich social lives, it is possible that they think of people with rich affective experiences. If that explanation holds, then clarifying the target of comparison should eliminate the overall tendency to overestimate others' greater preference for goods. Studies 9A–9B aims to clarify the target of comparison to investigate this possibility.

The second goal of Studies 9A–9B is to test an additional prediction based on intuitive confidence theory (Simmons & Nelson, 2006). When people form an initial intuition with great confidence, people still favor that intuition even after being directly informed that other information has rendered the intuition irrelevant. In the strongest example of this, Simmons and Nelson (2018) show that people make choices favoring an intuition even after they have personally established the standards to mitigate that initial judgment. For example, when people are asked whether a high performing company (like Amazon) will increase its stock price, most people say yes. Those same people are asked to estimate how much the price will rise, to give an exact representation of their intuition. Critically, when the participants are then asked whether the stock price will rise more than the amount that they just estimated, most people still say yes. Even when faced with explicitly equating information, a strong intuition still guides subsequent choice.

The same should apply in estimates of the experiences of others. People start with an initial intuition (e.g., "chocolates are good. I bet other people really like chocolates.") and do not adjust it sufficiently. Might the same tendency persist even when people are told that they are judging someone who is explicitly similar in valuation? Our theorizing predicts that when faced with a more difficult question (e.g., "would this other person be willing to pay more for the same amount of enjoyment?"), people may again favor their initial intuition and instead answer the easier question (e.g., "are other people, in general, willing to pay more than me?"). Because of this, we predict that despite being asked to consider another person who is explicitly similar in valuation, the tendency to overestimate others' valuations and preferences will persist, and that furthermore, it will expose an unusual phenomenon in perceived valuations.

## Method

**Participants and design.** Participants ( $N = 1,673$ ) recruited from Luth Research completed an online study for payment. Participants were randomly assigned to one of two different matched conditions: matched WTP or matched enjoyment. Only participants who answered the attention check correctly ( $N = 1,238$ ) were included in the analyses (preregistration: <https://aspredicted.org/bm79e.pdf>).

**Procedure.** Participants in the *matched WTP* conditions were first asked how much they would be willing to pay for a 12-piece box of Godiva Chocolatier Signature Chocolate Truffles. Next, they rated how much they would enjoy the box of chocolates on a 101-point scale (0 = *not at all* to 100 = *extremely*). Then, they were asked to consider another participant completing the survey who also indicated that they would be willing to pay the same amount for the box of chocolates that they indicated they would. Finally, they rated how much they thought that person would enjoy the box of chocolates on the same 101-point scale.

Participants in the *matched enjoyment* conditions were first asked how much they would enjoy a 12-piece box of Godiva Chocolatier Signature Chocolate Truffles on a 101-point scale (0 = *not at all* to 100 = *extremely*). Next, they indicated how much they would be willing to pay for the box of chocolates. Then, they were asked to consider another participant completing the survey who also indicated that they would enjoy the box of chocolates as much as they indicated they would. Finally, they indicated how much they thought that person would be willing to pay for the box of chocolates.

After completing these measures, participants completed a comprehension check. Those in the matched enjoyment condition were asked: "In this survey, we asked you to consider another participant. How much did we say the other participant indicated they would enjoy a 12-piece box of Godiva Chocolatier Signature Chocolate Truffles?" The choice options were "less than you would," "as much as you would," and "more than you would." Similarly, those in the matched WTP condition were asked: "In this survey, we asked you to consider another participant. How much did we say the other participant indicated they would pay for a 12-piece box of Godiva Chocolatier Signature Chocolate Truffles?" Again, the choice options were "less than you would," "as much as you would," and "more than you would." As previously mentioned, only those who correctly answered this comprehension check were included in the analyses below to ensure that any effects did not rely on inattention or misunderstanding on the part of the participant.

## Results and Discussion

Even when asked about another person who would enjoy the box of chocolates just as much as they would, participants believed that those others would be willing to pay significantly more for the box of chocolates than they themselves would ( $M_{\text{Difference Score}} = \$1.38$ ,  $SD_{\text{Difference Score}} = \$4.21$ ), one-sample  $t(593) = 7.96$ ,  $p < .001$ ,  $d = .33$ .

Conversely, even when asked about another person who would pay just as much for the box of chocolates as they would, participants believed that those others would enjoy the box of chocolates significantly more than they themselves would

( $M_{\text{Difference Score}} = 4.22$ ,  $SD_{\text{Difference Score}} = 16.31$ ), one-sample  $t(636) = 6.53$ ,  $p < .001$ ,  $d = .26$ .

As we mentioned in the introduction of this study, we had two main goals. The first was to identify whether our effects were operating through the idiosyncratic consideration of an unrepresentative other or whether, as we had theorized, the effect operated through influencing the perceived experience of others. Because this study forced people to consider the experience of someone whose evaluation was both known and familiar (i.e., it matched their own), the persistence of the effect would seem to challenge the possibility that people are considering an unrepresentative other.

The second goal of this study was to see whether the predictions of intuitive confidence were borne out in revealing a novel paradox in perceptions of others' valuations. People start with an automatic evaluation of the target (e.g., "people like chocolates.") which leads to a default judgment (e.g., "I bet they would pay more for it than I would."). When subsequent information is presented which could equate ratings of self and other (i.e., that the two are matched on anticipated enjoyment), the initial intuition still wins out, and people still seem to infer that another person would give a more extreme response.

These results are noteworthy. In practice, people pay for enjoyment, and some people are willing to pay more than others. If two people get the same enjoyment out of a box of chocolates, the person who is willing to pay more is effectively placing a higher monetary value on each unit of enjoyment. Those two chocolate consumers might be equated for their anticipated enjoyment, but one might be more concerned with getting a better deal (enjoyment per dollar) than the other. Considered from that perspective, we should rationally expect people to be consistent in their perceptions of others. They might perceive that someone else would be more frugal with their enjoyment consumption (a higher enjoyment per dollar) or more of a spendthrift (a lower enjoyment per dollar), but it would be irrational to simultaneously believe both.

People, however, do not seem to be rational in their beliefs. When matched with someone on enjoyment, people think that person would be willing to pay more for the same item (a perception that the other person gets less enjoyment per dollar). But when matched with someone on WTP, people think that person would get more enjoyment (a perception that the other person gets *more* enjoyment per dollar). People not only seem to have a belief that other people like things more than they do themselves, but they apply that belief in a way that entirely ignores the conditional relationship between the valuation of a good and the utility derived from it. That is, people do not simply think of an inaccurate average person in estimating the valuation of others. Instead, they also distort the scale of evaluation by focusing on one dimension of others' valuation but ignoring the other dimension of others' valuation that they are matched on. When evaluating someone identical on one measurement of liking, people thought other people would show more liking on an alternative measure. We replicated these results using different products (bonsai tree, a bottle of wine, and a box of pens) in [Supplemental Study S3](#). The details about this study can be found in the [online supplemental materials](#).

Study 9B tests whether this paradoxical result replicates with willingness-to-wait.

### Study 9B: Replication With Willingness-to-Wait

We previously demonstrated that people overestimate others' valuation and utility using three different valuation metrics: WTP, liking, and willingness-to-wait. In Study 9B, we test whether relying on core representations to estimate others' evaluations again leads to paradoxes about others' valuation and utility measured across those three metrics. In other words, we intended to investigate whether using the match paradigm in Study 9A with another measure of preference (wait time) also leads to paradoxical beliefs about others' preferences.<sup>8</sup>

### Method

**Participants and design.** Participants ( $N = 3,267$ ) recruited from MTurk completed an online study for payment. Participants were randomly assigned to one of four between-subjects conditions: (a) *enjoyment at matched wait time* ( $n = 617$ ); (b) *WTP at matched wait time* ( $n = 647$ ); (c) *wait time at matched enjoyment* ( $n = 791$ ); and (d) *wait time at matched WTP* ( $n = 727$ ). The product was a special early screening of *Star Wars: The Last Jedi*. Only participants who answered the attention check correctly ( $N = 2,782$ ) were included in the analyses (preregistration: <https://aspredicted.org/5re8z.pdf>).

**Procedure.** All participants made three evaluations: (a) how they would evaluate the movie screening on one measure, (b) how they would evaluate the movie screening on a second measure, and (c) how a similar other (one that matched them on the first measure) would evaluate the movie screening on the second measure on which they were not matched. For example, in the *enjoyment at matched wait time* condition, participants were asked how much they would be willing to wait to make sure they could attend an early screening of *Star Wars: The Last Jedi* (in minutes), then asked to evaluate how much they would enjoy attending an early screening of *Star Wars: The Last Jedi* on a 101-point scale (0 = *not at all* to 100 = *extremely*). Finally, they were asked to consider another participant taking the survey who reported being willing to wait the same amount of time as them for the early screening and asked how much they thought this other participant would enjoy the early screening on the same 101-point scale.

For conditions that included both wait time evaluations and WTP evaluations, we clarified that the evaluation was in lieu of the other evaluation. That is, prior to making their evaluation, participants were asked to consider "If instead of [paying/waiting in line] to make sure that you could attend that early screening, you would have to [wait in line/pay] to make sure that they could attend the early screening. . .".

### Results and Discussion

Conceptually replicating Study 9A, we demonstrated the value trade-off paradox for both wait time-enjoyment and wait time-WTP. Even when asked about another person who would be willing to wait the same amount of time for the early screening, participants believed that those others would enjoy the early screening significantly more ( $M = 59.28$ ,  $SD = 28.98$ ) than they themselves would ( $M = 56.69$ ,  $SD = 29.46$ ), one-sample  $t(616) = 4.79$ ,  $p < .001$ ,  $d = .19$ . At the same time, even when asked about another person who would enjoy attending the early screening just

as much as they would, participants believed that those others would wait significantly longer ( $M = 37.09$ ,  $SD = 44.40$ ) to attend the early screening than they themselves would ( $M = 24.69$ ,  $SD = 25.99$ ), one-sample  $t(790) = 10.27$ ,  $p < .001$ ,  $d = .36$ .

Similarly, even when asked about another person who would be willing to wait the same amount of time for the early screening, participants believed that those others would pay significantly more ( $M = \$16.73$ ,  $SD = \$20.33$ ) for the early screening than they themselves would ( $M = \$11.21$ ,  $SD = \$23.52$ ), one-sample  $t(646) = 6.26$ ,  $p < .001$ ,  $d = .25$ . At the same time, even when asked about another person who would be willing to pay the same amount for the early screening, participants believed that those others would wait significantly longer ( $M = 43.81$ ,  $SD = 56.50$ ) to attend the early screening than they themselves would ( $M = 31.79$ ,  $SD = 37.81$ ), one-sample  $t(726) = 10.47$ ,  $p < .001$ ,  $d = .39$ .

These results suggest that the observed bias is a broader phenomenon that consistently emerges across different metrics of value when people are making judgments about the valuation and preferences of others. This is a striking social judgment error. We hypothesize that this judgment error is specific to comparative judgments of valuation and preference. When considering the valuations of others, people have a strong general intuition: Other people value things more. On the other hand, the implied relationship of enjoyment to dollar or enjoyment to wait time does not come to mind intuitively, and so people end up with inconsistent responses.

Nevertheless, a more mundane possibility is that people have a more general shortcoming in evaluating the relationship between different estimates. That is, perhaps people are generally imperfect at perceiving stable ratios between *any* two correlated measures. By that reasoning, people are not just inconsistent in judgments of enjoyment-per-dollar, but also pounds-per-inches of height or kindness-per-donated dollar. In [Supplemental Studies S5 and S6](#), we examined whether or not judgments of other measures show similar irrationalities in different, nonpreference domains (height and weight; kindness and donation amounts). Contrary to this alternative explanation, in nonvalue domains, people's perceptions were consistent across different manners of elicitation. This seems to demonstrate that for other domains, people are internally rational. In other words, it is not that people are purely incapable of consistently answering questions using this matched procedure. Instead, there appears to be something unusual about perceptions in the domains of value. Additional details about these studies can be found in the [online supplemental materials](#).

### Study 10: Distribution of Others' Payments and Enjoyment

We have provided robust evidence that people have persistent beliefs that others have more intense experiences than they themselves do. We argue that, when estimating the valuations and preferences of others, people rely on the core representation of given products and experiences. One potential way to disrupt this

<sup>8</sup> We initially conducted a study ( $N = 1,188$ ) examining this question without preregistering and using a smaller sample (see [Study S4](#) in the online supplemental studies and materials). We conducted Study 9B to confirm the results of this study with a larger sample.



reliance on core representations is to explicitly ask people to consider a wider array of valuations and preferences. That is, by forcing people to consider those whose valuations are atypical (e.g., people who might dislike a generally positively regarded experience), they will be required to think about the heterogeneity of the experience itself. A typical Rio De Janeiro vacationer brings to mind all the positive features of the city, but thinking of a less-positive Rio De Janeiro vacationer might remind people of exactly those less salient features which are not necessarily so positive. Study 10 tested whether asking people to estimate a distribution of others' preferences (thereby more explicitly including those who might dislike a positive product), rather than a point estimate of others' preferences, attenuates the overestimation bias.

## Method

**Participants and design.** Participants ( $N = 1,964$ ) recruited by Luth Research completed an online study for payment. Participants were randomly assigned to a 2 (person: self or other)  $\times$  2 (order: self first or other first)  $\times$  2 (DV: WTP or liking)  $\times$  2 (format: point estimate or distribution) mixed-factors design with person as the only within-subjects factor. Only participants who answered the attention check correctly ( $N = 1,705$ ) were included in the analyses (preregistration: <https://aspredicted.org/wc5f7.pdf>).

**Procedure.** Participants evaluated a card for any two drinks at Starbucks that they were told had a maximum value of \$10. Participants reported their WTP or liking for this card for themselves as well as other participants taking the same survey. In the WTP conditions, participants reported WTP for this card. In the liking conditions, they reported liking for this card. The order of self and other measures was randomized.

For self-evaluations, participants reported their WTP on a scale that ranged from \$0 to \$10 with \$1 intervals in the WTP conditions or their liking on an 11-point scale (0 = *not at all* to 10 = *extremely*) in the liking conditions. For other evaluations, participants evaluated the Starbucks card for others taking the same survey in one of two different formats: (a) point estimates or (b) distributions. In the *point estimate* conditions, participants estimated others' WTP by selecting an amount from the same \$0 to \$10 scale as self evaluations (WTP conditions) or estimated others' liking on the same 11-point scale as self evaluations (liking conditions). In the *distribution* conditions, participants were asked to allocate percentage of other survey takers who would choose each of the 11 different choice options. That is, participants were asked to allocate the percentage of others who would indicate that they would pay each of the \$0 to \$10 payment options (WTP conditions) or the percentage of others who would indicate how much they would like the product on each of the 0 to 10 rating options (liking conditions). The total percentage was required to add up to 100%.

## Results and Discussion

**WTP.** To compare the distribution condition to the point estimate condition, we first calculated the average WTP of others in the distribution condition by multiplying the allocated percentages to each dollar amount by the dollar amount. We then submitted WTP responses to a 2 (person: self or other)  $\times$  2 (order: self first or other first)  $\times$  2 (format: point estimate or distribution) mixed ANOVA with person as the only within-subjects factor.

We observed a significant main effect of person,  $F(1, 885) = 74.31$ ,  $p < .001$ ,  $\eta_p^2 = .08$ , such that participants believed that others would pay more for a card for any two drinks at Starbucks ( $M = \$6.00$ ,  $SD = \$2.79$ ) than they would ( $M = \$5.09$ ,  $SD = \$3.36$ ). More importantly, we wanted to see whether the overestimation bias we observed in the previous studies would be attenuated in the distribution conditions. First, before assessing the influence of the manipulation, it is worthwhile to consider how people perceive the distribution of others' evaluations. Figure 4 shows the distribution of values for the self (in blue) and the perceptions of others (in red). The self-ratings show a hint of central tendency, but are largely dominated by people using the lowest value on the scale (\$0), the midpoint (\$5), and the highest value (\$10). Although this is hardly an intuitive distribution (to these researchers, at least), participants were impressively accurate in capturing it. The fidelity of their representation gives further weight to the idea that people are not selectively unable to perceive the full distribution of responses, but rather than they are somehow distorting that perception when making their overall judgments.

Knowing that participants were generally good at predicting the distribution of responses, the influence of the manipulation was striking. Is the overestimation effect attenuated after people are asked to recall the distribution of valuation? Consistent with this possibility, the Person  $\times$  Format interaction was significant,  $F(1, 885) = 42.99$ ,  $p < .001$ ,  $\eta_p^2 = .05$ . When estimating others with a point estimate, people believed others would pay more for the Starbucks card ( $M = \$6.67$ ,  $SD = \$2.51$ ) than they would ( $M = \$5.13$ ,  $SD = \$3.33$ ), one-sample  $t(481) = 11.00$ ,  $p < .001$ ,  $d = .50$ . However, when estimating others with a distribution, people believed others would pay about as much for the Starbucks card ( $M = \$5.21$ ,  $SD = \$2.91$ ) as they would ( $M = \$5.04$ ,  $SD = \$3.39$ ), one-sample  $t(406) = 1.04$ ,  $p = .30$ ,  $d = .05$  (see Figure 4). This attenuation did not significantly change based on which person they evaluated first (self or other),  $F(1, 885) = 2.11$ ,  $p = .147$ ,  $\eta_p^2 = .002$ .

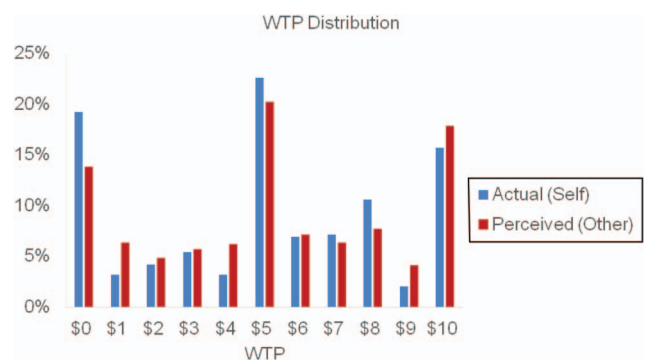
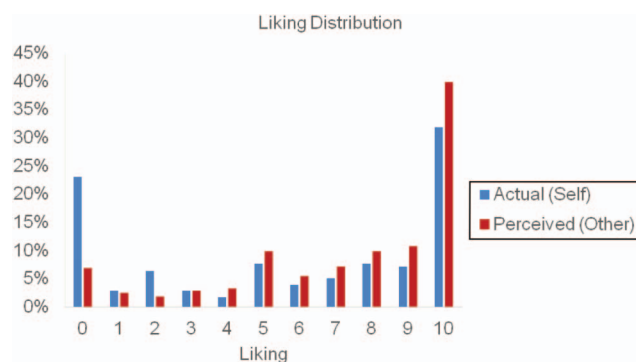


Figure 4. Actual and estimated percentage of participants who would pay a given price (on X-axis) for a card that buys two drinks at Starbucks in the WTP distribution condition in Study 10. Actual percentage was calculated as the proportion of participants who indicated that they would pay the specific price out of all the participants in the WTP distribution condition. The estimated percentage was calculated as the average percentage estimated for each price point (e.g., for \$0, the average of all participants' guesses of the percentage of others who would indicate that they would pay \$0). WTP = willingness-to-pay. See the online article for the color version of this figure.

**Liking.** As with WTP, to compare the distribution condition to the point estimate condition, we first calculated the average liking of others for a card in the distribution condition by multiplying the allocated percentages to each liking rating option by the rating. We then submitted WTP responses to a 2 (person: self or other)  $\times$  2 (order: self first or other first)  $\times$  2 (format: point estimate or distribution) mixed ANOVA with person as the only within-subjects factor. We observed a significant main effect of person,  $F(1, 812) = 251.01, p < .001, \eta_p^2 = .31$ , such that participants believed that others would like a card more ( $M = 7.46, SD = 2.33$ ) than they would ( $M = 5.44, SD = 3.98$ ). More importantly, the Person  $\times$  Format interaction was significant,  $F(1, 812) = 10.06, p = .002, \eta_p^2 = .01$ . When estimating others with a point estimate, people believed others would like the Starbucks card more ( $M = 7.59, SD = 2.22$ ) than they would ( $M = 5.17, SD = 3.89$ ), one-sample  $t(406) = 14.20, p < .001, d = .70$ . When estimating others with a distribution, this effect was significantly attenuated but still significant: People still believed others would like the Starbucks card more ( $M = 7.33, SD = 2.43$ ) as they would ( $M = 5.71, SD = 4.06$ ), one-sample  $t(408) = 8.54, p < .001, d = .42$  (see Figure 5). Therefore, while considering the entire distribution of others' liking did not entirely eliminate overestimation, it did reduce it significantly. Moreover, this attenuation did not significantly change based on which person they evaluated first (self or other),  $F(1, 812) = .97, p = .326, \eta_p^2 = .001$ .

Thus, for both WTP and enjoyment, people's overestimation of others' valuations was attenuated when considering the full distribution of others' possible valuation and liking. These results are consistent with the explanation that the overestimation bias in valuation stems from people's reliance on core representations of experiences. When this reliance is disrupted, the overestimation bias significantly diminishes.<sup>9</sup>



**Figure 5.** Actual and estimated percentage of participants who would like a card that buys two drinks at Starbucks a specific amount in the liking distribution condition in Study 10. Actual percentage was calculated as the proportion of participants who indicated that they would like the card a specific amount out of all the participants in the liking distribution condition. The estimated percentage was calculated as the average percentage estimated for each liking point (e.g., for a 0 rating, the average of all participants' guesses of the percentage of others who would indicate that they would like the card a 0). See the online article for the color version of this figure.

## General Discussion

People are sometimes called upon to assess the preferences of others, assessments which we find to be prone to persistent biases. Across several studies, we find that across various measures of valuation and utility (i.e., WTP, enjoyment, and willingness-to-wait), people believe that others have more intense experiences than they themselves do (Studies 1–8). We propose that this overestimation of others stems from a narrow focus on the primary dimension of the option being evaluated (e.g., a trip to Rio De Janeiro is generally thought to be positive, whereas shaving your head is generally thought to be negative). But this only involves estimations of others. Self-assessments are further informed by the subtle vagaries of personal preferences, reducing the total preference intensity (e.g., Rio De Janeiro is encumbered by its hard-to-pronounce local language, and a shaved head is buoyed by the opportunity it affords for a novel scalp tattoo). Thus, personal evaluations are more moderate than are the estimates of the evaluations of others.

This intuition is strong enough that it is applied even when the target of comparison is explicitly similar to the self (Studies 9A–9B). When asked to evaluate someone with an identical WTP, people think that person will anticipate more enjoyment; and when evaluating someone with identical anticipated enjoyment, people think that person will have a higher WTP. In combination, people can demonstrate the paradoxical belief that others are willing to pay more for the same level of enjoyment (when asked about someone identical in enjoyment) or that others are willing to pay less for the same level of enjoyment (when asked about someone identical in WTP). Finally, explicitly prompting people to think about the full distribution of others' possible valuations significantly interrupted the intuitive process of overestimation based on the core representation of objects being considered (Study 10).

## Relation to Previous Research

Why do people have such persistent judgmental errors when estimating the evaluations of others? After all, people are not blind to the evaluations of others. People frequently observe the choices of others, and at least occasionally, are told something about the preferences which led to those choices. Research indicates that judgments about values and preferences are often inherently automatic (Chaiken & Trope, 1999; Kahneman & Frederick, 2002; Kahneman, 2003; Simmons & Nelson, 2006, 2018; Sloman, 1996). Understanding the trade-offs between evaluative metrics (e.g., a longer wait vs. a higher price), however, is more complicated (e.g., Tversky, Sattath, & Slovic, 1988). When reporting their own evaluations, people have the benefits of each metric being accessible and generally reliable, and consequently, evaluative trade-offs are more likely to be consistent. When predicting the evaluations of others, on the other hand, people do not have the same basis of knowledge. Without knowledge of how other people trade off between evaluative metrics, people appear to ignore them

<sup>9</sup> As this was our initial test of this distribution manipulation, we conducted a replication (Supplemental Study S7) with a new stimulus: a chocolate bar ( $N = 1,002$ ). We generally replicated the results: Considering a distribution of others' preferences and valuations attenuated the overestimation bias (see online supplemental materials for additional details).

altogether. Consequently, they use intuitive but incomplete heuristics that people experience things more intensely, which can be misapplied in the case of similar others (i.e., those who would like a good as much as they would or would pay as much for a good than they would).

Previous research in judgment and decision-making documents abundant evidence that people do not always hold stable preferences but construct them on the spot when they are making decisions (Bettman, Luce, & Payne, 1998; Fischhoff, 2013; Payne, Bettman, & Johnson, 1992; Slovic, 1995). If preferences are partially constructed for the self, they might be entirely constructed when judging others. Studies 9A–9B demonstrate that while people's own valuation of a good remains stable, their beliefs about others' valuation of the same good reverse depending on how they are asked about others' valuation. More specifically, people believed that others derived simultaneously more and less utility from the same goods than they did.

The overestimation bias we document also offers a new approach to understanding the endowment effect and why selling prices tend to exceed buying prices. Previous research has largely focused on a "pain-of-losing" account for this phenomenon, which proposes that people feel significantly more pain when selling their good than others feel when acquiring the same good (Kahneman & Tversky, 1982; Thaler, 1980; Tversky & Kahneman, 1991). Another explanation more recently put forth by Weaver and Frederick (2012) hypothesizes that instead sellers and buyers use different reference prices. Sellers typically focus on market prices in determining their selling price, whereas buyers typically focus on their own valuation. Because market prices tend to be higher than people's valuations (Kahneman, Knetsch, & Thaler, 1991) and both parties are averse to bad deals, selling prices tend to exceed buying prices. Our overestimation bias account suggests that in addition to these explanations, people's expectation that others derive more value from goods might also contribute to a discrepancy in buying and selling prices. In particular, sellers may believe that buyers would value the good more than they themselves would, leading them to set higher selling prices.

### Alternative Accounts and Directions for Future Research

This article reports 12 experiments documenting the existence, robustness, and consequence of the overestimation bias. We also conducted a handful of additional investigations to try to understand the forces that may moderate our effects. Though these studies do not authoritatively answer why people overestimate others' valuation, in combination they may provide some hints. We review two of those investigations, and report them in further detail in the [online supplemental materials](#).

**Others with extreme preferences.** Study 9A introduced the matching paradigm as a strong tactic for controlling how people generate an exemplar when estimating the evaluation of others. An alternative approach, we thought, might be to simply heighten the salience of some comparison others who are more or less positive about the same stimulus. If people are spontaneously thinking of an enthusiastic consumer, then forcing them to consider the behavior of an unenthusiastic consumer might change their estimate. We examined this exceptional other account in two additional studies described in detail in the [online supplemental materials](#).

First, in [Supplemental Study S8](#) ( $N = 807$ ), we recruited people who self-identified as having extreme preferences to investigate whether the overestimation bias would persist. Specifically, we recruited self-identified fans of *Star Wars* movies and asked them to estimate either: (a) the average *Star Wars* fan's or (b) the average U.S. person's evaluations of a *Star Wars* product. Though these *Star Wars* fans rationally understood that the average U.S. person's evaluation of a *Star Wars* product would be less extreme than their own, their overestimation emerged when considering the average *Star Wars* fan, assuming that the average *Star Wars* fan would evaluate the product more positively than they themselves would.

Second, in [Supplemental Study S9](#) ( $N = 1,214$ ), we used the match paradigm in Study 9A with an additional factor. In addition to examining how people view *identical* others (i.e., those matched on either enjoyment or WTP), we explored people's estimates for either: (a) a person who had greater preference for the product (i.e., would pay \$5 more than they would for the product [higher WTP other] or would enjoy the product five units more than they would [higher enjoyment other]), or (b) a person who had lesser preference for the product (i.e., would pay \$5 less than they would for the product [lower WTP other] or would enjoy the product five units less than they would [lower enjoyment other]). By our reasoning, it is possible that explicitly considering a less enthusiastic consumer would disrupt people's intuitions for their preferences, thereby eliminating overestimation. We first replicated the paradoxical results of Study 9A when people considered identical others: People assumed both that those matched on enjoyment would pay more for the product than they would, but also that those matched on WTP would enjoy the product more than they would. But importantly, people asked to consider *lower enjoyment others* (i.e., those who would enjoy the product five units less than they would) rationally assumed that those others would pay less for the product than they would, and people asked to consider *lower WTP others* (i.e., those who would pay \$5 less than they would for the product) rationally assumed that those others would enjoy the product less than they would.

Together, the results from these two supplemental studies bolster our finding in Studies 9A–9B that the bias cannot be fully explained by the salience of others with extreme preferences or the extremity of one's own preferences. When people explicitly consider others who are less positive toward a product, people display rational responses. However, when considering average others or those who should have similar preferences, the overestimation bias persists.

Although people often inaccurately predict others' preferences, they are more likely to be accurate about the relative difference between their own and others' preferences of certain experiences. For instance, if a parent were asked how much they and others would like their child's drawing, they would no doubt recognize that their liking of the drawing would be greater than that of others. Or if people are explicitly told that someone likes a product *less* than they themselves do, this also appears to disrupt the reliance on intuitive core representations. The results in Study 10 are consistent with this logic: People can more accurately predict others' preferences when they are explicitly prompted to consider others whose preferences are not consistent with the core representation of a stimulus. Therefore, people are capable of understanding others' preferences, but they do not spontaneously consider and integrate the entire distribution of possible preferences unless they are explicitly compelled to do so.



The combination of the above points does highlight an interesting parallel account, one that we can articulate with some clarity, but one that our present findings can neither perfectly rule out nor perfectly rule in. In our theorizing, the prospective visitor to Rio De Janeiro forecasts a positive experience encumbered by a small number of idiosyncratic negative experiences. That person, when judging others, starts with the core representation of the experience (i.e., that a visit to Rio De Janeiro is enjoyable), and that confidence in that initial intuition means that they do not adjust from there. Accordingly, whereas personal assessments of Rio De Janeiro are somewhat middling, others are perceived to be more positive. The alternative account focuses not on the mixture of experience within an individual that moves a high rating to a lower rating, but rather the mixture of experiences across people that produces many positive evaluations, but also some idiosyncratically individually low ratings for generally positive stimuli. Consider again the person evaluating the trip to Rio De Janeiro. On average, that person is probably positive (say, an 85 on a 101-point scale), but some people might be quite negative (perhaps they are actively avoiding irritating in-laws back in Brazil), and give an extremely low evaluation of the potential visit. The average person and the negative person both have equal weight in the overall true average, but they might not have equal weight in how people form their perceptions of the average other. In essence, it may be the case that people are fully capable of integrating both the core representation of a prospect with the more unusual negative features; they accurately recognize that most people think that Ipanema beach is beautiful, and they accurately recognize that most people are nevertheless bothered by the risk for potential theft, but they fail to capture that for some people the latter factor is so significant that it overwhelms the former. That is, they accurately perceive the experiences of others, but they do not consider all of those experiences when estimating the average experience.<sup>10</sup>

There is merit to this account. First, even for very positive stimuli, there are always a number of participants whose valuations are quite low. Consider, just as an example, the density plot for WTP for a movie ticket from Study 1 (see Figure 1A). The mean is not low, but the distribution is hardly normal, and a sizable fraction of participants (26.4%) say that they are willing to pay \$0. Perhaps it is exactly that segment of the population that people are failing to identify when constructing their averages. The distributions generated by participants in Study 10 partially challenge the extreme version of that possibility. Participants generated very accurate representations of other people's WTP (albeit less accurate for other people's enjoyment), but still showed the overall bias. Still, it may be the case that people are capable of bringing the full and accurate representation of the distribution to mind, but they do not do so unless prompted.

For now, we remain agnostic. It could be the case that the core representation of a positive product or experience creates an intuition that biases predictions about everyone upward, or it may be the case that creates a mental sample that is biased by selecting out those people who are uncharacteristically negative. Some of our data seems more consistent with one, but none of it is so consistent to eliminate the possibility of the other. We think that future research can hopefully untangle those (and we hope that we are the researchers who do so).

**Specific or more familiar others.** This research documents a robust bias in comparative judgments in valuation and preferences that extends to a wide range of stimuli. As we demonstrate, this bias persists even in many situations that might be expected to attenuate it. Nevertheless, the overestimation bias is unlikely to persist in all social predictions. For example, when people think of someone familiar (e.g., a family member or a close friend) and that person's specific preferences for Rio De Janeiro, they would presumably be more likely to draw upon the preference nuances they know, rather than blindly rely on the intuitive core representation of Rio De Janeiro. Therefore, it seems reasonable to predict that, as the other becomes more familiar, the overestimation bias could attenuate or even reverse.

**Underestimation bias.** We interpret the documented bias as people perceiving others having more intense experiences than they do. As aforementioned, however, people also underestimate the intensity of experiences of others who are in a different emotional state (Van Boven et al., 2000; Van Boven & Loewenstein, 2003). Although it is not the main focus of this research, Study 6 considered some negative experiences that could reasonably be assumed to induce embarrassment (e.g., composing and reciting a poem about love and walking around with a t-shirt that says "Repent!"). Somewhat contrary to an empathy gap account, participants in Study 6 predicted that others would be willing to pay *more* to avoid embarrassing experiences.

The observed discrepancy in Van Boven, Loewenstein, and Dunning (2005) and the results in Study 6 may be at least partially attributable to the difference between the measure we used—WTP to avoid—and the one used by Van Boven et al. (2005)—compensation demanded. Although conceptually nearly identical, the two measures have been shown to vary substantially and operate inconsistently (Hammack & Brown Jr., 1974; Harless, 1989; cf., Knetsch & Sinden, 1984; Marshall, Knetsch, & Sinden, 1986; Plott & Zeiler, 2005; Rowe, D'Arge, & Brookshire, 1980). To probe more directly whether the discrepancy in Van Boven et al. (2005) and this research is driven by how preferences are measured, we conducted a replication of Study 6, but this time we also varied the valuation measure such that some participants indicated the minimum compensation required, whereas the rest answered the maximum WTP to avoid negative experiences (See Supplemental Study S10 for details). As we document in this article, people thought that others would set a higher maximum WTP to avoid the negative experience. However, when we instead use the "minimum compensation required" measure, the effect reverses: People think that others will require less compensation than they would themselves. Given the various accounts behind the WTP and WTA gap the previous literature identified, one or more of those accounts could be relevant for the observed reversal of valuation for self and other. Future research could more directly examine this inconsistency in measurements of valuation judgments and the source of the reversal of social predictions with two supposedly equivalent economic metrics.

**Retrospective evaluations.** In all studies we report, participants indicated their own valuation and preferences and making predictions about others *before* actually consuming experiences. It

<sup>10</sup> This alternative was pointed out by an attentive reviewer. We are authentically grateful to that person for identifying it.



is possible that the reported overestimation bias is specific to prospective judgments, or the lack of actual consumption, and does not appear with retrospective judgments after consuming experiences. We tested this possibility in [Supplemental Study S11](#). In this study, participants evaluated a short film clip either before watching it or after having just watched it. If overestimation of others' valuations occurs only when people are merely thinking about consuming an experience, our effect should be eliminated when people have just consumed the experience. Contrary to this prediction, people displayed overestimation of others' liking, regardless of whether they had or had not actually consumed the experience. Additional details of this study can be found in the [online supplemental materials](#).

**Different types of products and experiences.** In our present investigations, we found overestimation across a variety of products and experiences. As we previously mentioned, our effect sizes varied based on the products and experiences we tested. One possibly fruitful direction for future research is to investigate which types of products and experiences are more or less likely to demonstrate the overestimation we see here. As an initial investigation (reported in [Supplemental Study S12](#)) prompted by a suggestion from a reviewer, we tested three products whose value was partially tied to what they signaled to other people. It is possible that social signaling products, in part because they intrinsically highlight heterogeneity of preferences, may show an attenuated or reversed overestimation effect. That is, if a person believes that a product makes them appear morally or intellectually superior, that person might indicate that they enjoy and are willing to pay more for this product compared with others as a way of signaling moral and intellectual superiority to themselves and others. Participants reported their WTP for and liking/disliking on a bipolar scale for: (a) a Make America Great Again (MAGA) hat (a political product likely to be perceived as positive for some and negative for others); (b) a pack of metal straws (a product that may signal whether the person is environmentally conscious); and (c) a book of *New Yorker* cartoons (a product that may signal intellectualism) for themselves and for the other participants in the survey. For all products, participants believed others would be willing to pay more for these products than they themselves would. The bipolar liking/disliking measures suggested a slightly more ambiguous response. For the MAGA hat, we found that participants strongly expected that they would dislike the hat more than others (possibly because the university participant pool was likely liberal). For the metal straws, we found (nonsignificant) underestimation. Finally, for the *New Yorker* cartoons, we found no difference between self and other. The metal straws and *New Yorker* cartoons seem to suggest that the overestimation bias may be attenuated for products that are tied to social signals. Furthermore, the reversal with MAGA hats may suggest that in cases when the core representation of a product is complicated (some people hate the product but others like it), the overestimation bias is likewise complicated.

We do not have a perfect explanation for this pattern of results, but we can speculate. When evaluating a vacation to Rio De Janeiro, people can safely assume that the general valence is positive for everyone (it is a vacation after all), but then may neglect how attuned other people will be to potential idiosyncratic shortcomings of the trip. Relative to a vacation, the MAGA hat may represent much more severe idiosyncrasies of preference; not only do people vary in how much they would like wearing it, many

people will actively dislike it, and some of those will dislike it a lot. Therefore, this complex distribution of others' preferences makes predictions of others' preferences considerably more difficult, and as is apparent from the data, people systematically err in answering this question. We think that further exploration of such complex stimuli would be a worthwhile endeavor.

## Implications

In social judgments, we are frequently called upon to make predictions about the evaluations of others, such as how much a friend will enjoy a recommended novel, how long a coworker will be willing to wait for useful feedback, or how much a potential buyer will be willing to offer for a used set of golf clubs. The overestimation bias has important implications on real-world economic and social decisions. Furthermore, the paradox we document suggests that someone who perfectly understands others' enjoyment for a good may nevertheless be imperfect at setting prices, simply because they fail to recognize how people trade off enjoyment against other valuation metrics, such as WTP and willingness-to-wait. This suggests that social and consumer judgments may suffer not only from general overestimation but also from an additional imperfect understanding of how people weigh trade-offs.

Many preferences are developed from a complex weighting of positives and negatives, the balance of which produces a summary evaluation for each individual. When judging the preferences of others, however, that complexity may be ignored and the weight of some attributes will seemingly tip the scales in the direction of a simpler, more intense preference.

## References

- Barasz, K., Kim, T., & John, L. K. (2016). The role of (dis)similarity in (mis)predicting others' preferences. *Journal of Marketing Research*, 53, 597–607. <http://dx.doi.org/10.1509/jmr.15.0226>
- Bettman, J. R., Luce, M. F., & Payne, J. W. (1998). Constructive consumer choice processes. *The Journal of Consumer Research*, 25, 187–217. <http://dx.doi.org/10.1086/209535>
- Chaiken, S., & Trope, Y. (1999). *Dual-process theories in social psychology*. New York, NY: Guilford Press.
- Deri, S., Davidai, S., & Gilovich, T. (2017). Home alone: Why people believe others' social lives are richer than their own. *Journal of Personality and Social Psychology*, 113, 858–877. <http://dx.doi.org/10.1037/pspa0000105>
- Epley, N., & Dunning, D. (2000). Feeling "holier than thou": Are self-serving assessments produced by errors in self- or social prediction? *Journal of Personality and Social Psychology*, 79, 861–875. <http://dx.doi.org/10.1037/0022-3514.79.6.861>
- Epley, N., Keysar, B., Van Boven, L., & Gilovich, T. (2004). Perspective taking as egocentric anchoring and adjustment. *Journal of Personality and Social Psychology*, 87, 327–339. <http://dx.doi.org/10.1037/0022-3514.87.3.327>
- Fischhoff, B. (2013). Value elicitation: Is there anything in there? *Judgment and Decision Making*, 46, 207–231.
- Frederick, S. (2012). Overestimating others' willingness to pay. *The Journal of Consumer Research*, 39, 1–21. <http://dx.doi.org/10.1086/662060>
- Hamack, J., & Brown, G. M., Jr. (1974). *Waterfowl and wetlands: Toward bioeconomic analysis* (1st ed.). London, UK: Routledge. <http://dx.doi.org/10.4324/9781315659039>
- Harless, D. W. (1989). More laboratory evidence on the disparity between willingness to pay and compensation demanded. *Journal of Economic Behavior & Organization*, 11, 359–379. [http://dx.doi.org/10.1016/0167-2681\(89\)90035-8](http://dx.doi.org/10.1016/0167-2681(89)90035-8)

- Jung, M. H., Nelson, L. D., Gneezy, A., & Gneezy, U. (2014). Paying more when paying for others. *Journal of Personality and Social Psychology*, 107, 414–431. <http://dx.doi.org/10.1037/a0037345>
- Kahneman, D. (2003). A psychological perspective on economics. *American Economic Review*, 93, 162–168. <http://dx.doi.org/10.1257/000282803321946985>
- Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. *Heuristics and biases: The Psychology of Intuitive Judgment*, 49, 81.
- Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1991). Anomalies: The endowment effect, loss aversion, and status quo bias. *The Journal of Economic Perspectives*, 5, 193–206. <http://dx.doi.org/10.1257/jep.5.1.193>
- Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, 3, 430–454. [http://dx.doi.org/10.1016/0010-0285\(72\)90016-3](http://dx.doi.org/10.1016/0010-0285(72)90016-3)
- Kahneman, D., & Tversky, A. (1982). Variants of uncertainty. *Cognition*, 11, 143–157.
- Katz, D., & Allport, G. (1931). *Student attitudes: A report of the Syracuse University research study*. Syracuse, NY: Craftsman.
- Knetsch, J. L., & Sinden, J. A. (1984). Willingness to pay and compensation demanded: Experimental evidence of an unexpected disparity in measures of value. *The Quarterly Journal of Economics*, 99, 507–521. <http://dx.doi.org/10.2307/1885962>
- Krueger, J. (1998). On the perception of social consensus. *Advances in Experimental Social Psychology*, 30, 163–240. [http://dx.doi.org/10.1016/S0065-2601\(08\)60384-6](http://dx.doi.org/10.1016/S0065-2601(08)60384-6)
- Krueger, J. (2000). The projective perception of the social world: A building block of social comparison processes. In J. Suls & L. Wheeler (Eds.), *The Plenum series in social/clinical psychology. Handbook of social comparison: Theory and research* (pp. 323–351). Dordrecht, the Netherlands: Kluwer Academic Publishers. [http://dx.doi.org/10.1007/978-1-4615-4237-7\\_16](http://dx.doi.org/10.1007/978-1-4615-4237-7_16)
- Kurt, D., & Inman, J. J. (2013). Mispredicting others' valuations: Self-other difference in the context of endowment. *The Journal of Consumer Research*, 40, 78–89. <http://dx.doi.org/10.1086/668888>
- Larsen, J. T., McGraw, A. P., & Cacioppo, J. T. (2001). Can people feel happy and sad at the same time? *Journal of Personality and Social Psychology*, 81, 684–696. <http://dx.doi.org/10.1037/0022-3514.81.4.684>
- Marshall, J. D., Knetsch, J. L., & Sinden, J. A. (1986). Agents' evaluations and the disparity in measures of economic loss. *Journal of Economic Behavior & Organization*, 7, 115–127. [http://dx.doi.org/10.1016/0167-2681\(86\)90001-6](http://dx.doi.org/10.1016/0167-2681(86)90001-6)
- Miller, D. T., & McFarland, C. (1987). Pluralistic ignorance: When similarity is interpreted as dissimilarity. *Journal of Personality and Social Psychology*, 53, 298–305. <http://dx.doi.org/10.1037/0022-3514.53.2.298>
- Moon, A., & Nelson, L. D. (in press). The uncertain value of uncertainty: When consumers are unwilling to pay for what they like. *Management Science*. <http://dx.doi.org/10.1287/mnsc.2019.3426>
- Nordgren, L. F., Banas, K., & MacDonald, G. (2011). Empathy gaps for social pain: Why people underestimate the pain of social suffering. *Journal of Personality and Social Psychology*, 100, 120–128. <http://dx.doi.org/10.1037/a0020938>
- Payne, J., Bettman, J. R., & Johnson, E. J. (1992). Behavioral decision research: A constructive processing perspective. *Annual Review of Psychology*, 43, 87–131. <http://dx.doi.org/10.1146/annurev.ps.43.020192.000511>
- Plott, C. R., & Zeiler, K. (2005). The willingness to pay-willingness to accept gap, the “endowment effect,” subject misconceptions, and experimental procedures for eliciting valuations. *The American Economic Review*, 95, 530–545. <http://dx.doi.org/10.1257/0002828054201387>
- Polman, E., Effron, D. A., & Thomas, M. R. (2017). Other people's money: Money's perceived purchasing power is smaller for others than for the self. *The Journal of Consumer Research*, 45, 109–125. <http://dx.doi.org/10.1093/jcr/tucx119>
- Rosch, E. (1975). Cognitive reference points. *Cognitive Psychology*, 7, 532–547. [http://dx.doi.org/10.1016/0010-0285\(75\)90021-3](http://dx.doi.org/10.1016/0010-0285(75)90021-3)
- Ross, L., Greene, D., & House, P. (1977). The “false consensus effect”: An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology*, 13, 279–301. [http://dx.doi.org/10.1016/0022-1031\(77\)90049-X](http://dx.doi.org/10.1016/0022-1031(77)90049-X)
- Rowe, R. D., D'Arge, R. C., & Brookshire, D. S. (1980). An experiment on the economic value of visibility. *Journal of Environmental Economics and Management*, 7, 1–19. [http://dx.doi.org/10.1016/0095-0696\(80\)90018-2](http://dx.doi.org/10.1016/0095-0696(80)90018-2)
- Simmons, J. P., & Nelson, L. D. (2006). Intuitive confidence: Choosing between intuitive and nonintuitive alternatives. *Journal of Experimental Psychology: General*, 135, 409–428. <http://dx.doi.org/10.1037/0096-3445.135.3.409>
- Simmons, J. P., & Nelson, L. D. (2018). *Six biases that are all the same: An introduction to comparative choice theory*. Unpublished Manuscript, Wharton School of Business, University of Pennsylvania, Philadelphia, PA.
- Simmons, J. P., Nelson, L. D., Simonsohn, U., & Life after P-Hacking. (2013). *Meeting of the society for personality and social psychology*, New Orleans, LA, 17–19 January 2013. Retrieved from <https://ssrn.com/abstract=2205186> or <http://dx.doi.org/10.2139/ssrn.2205186>
- Sloman, S. A. (1996). The Probative value of simultaneous contradictory belief: Reply to Gigerenzer and Regier (1996). *Psychological Bulletin*, 119, 27–30. <http://dx.doi.org/10.1037/0033-2909.119.1.27>
- Slovic, P. (1995). The construction of preference. *American Psychologist*, 50, 364–371. <http://dx.doi.org/10.1037/0003-066X.50.5.364>
- Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization*, 1, 39–60. [http://dx.doi.org/10.1016/0167-2681\(80\)90051-7](http://dx.doi.org/10.1016/0167-2681(80)90051-7)
- Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5, 207–232. [http://dx.doi.org/10.1016/0010-0285\(73\)90033-9](http://dx.doi.org/10.1016/0010-0285(73)90033-9)
- Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211, 453–458. <http://dx.doi.org/10.1126/science.7455683>
- Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *The Quarterly Journal of Economics*, 106, 1039–1061. <http://dx.doi.org/10.2307/2937956>
- Tversky, A., Sattath, S., & Slovic, P. (1988). Contingent weighting in judgment and choice. *Psychological Review*, 95, 371–384. <http://dx.doi.org/10.1037/0033-295X.95.3.371>
- Van Boven, L., Dunning, D., & Loewenstein, G. (2000). Egocentric empathy gaps between owners and buyers: Misperceptions of the endowment effect. *Journal of Personality and Social Psychology*, 79, 66–76. <http://dx.doi.org/10.1037/0022-3514.79.1.66>
- Van Boven, L., & Loewenstein, G. (2003). Social projection of transient drive states. *Personality and Social Psychology Bulletin*, 29, 1159–1168. <http://dx.doi.org/10.1177/0146167203254597>
- Van Boven, L., Loewenstein, G., & Dunning, D. (2005). The illusion of courage in social predictions: Underestimating the impact of fear of embarrassment on other people. *Organizational Behavior and Human Decision Processes*, 96, 130–141. <http://dx.doi.org/10.1016/j.obhdp.2004.12.001>
- Weaver, R., & Frederick, S. (2012). A reference price theory of the endowment effect. *Journal of Marketing Research*, 49, 696–707. <http://dx.doi.org/10.1509/jmr.09.0103>
- Ziano, I., & Villanova, D. (2019). *You'd use it more than me: Overestimating products' usefulness to others because of self-serving materialism attributions*. Retrieved from <https://psyarxiv.com/938m7/>

Received February 4, 2019

Revision received September 13, 2019

Accepted September 13, 2019 ■