Overestimating the Valuations of Others: People Perceive Others as Experiencing Everything More Intensely

Minah H. Jung,¹ Alice Moon², Leif D. Nelson³

¹New York University, ²University of Pennsylvania, ³University of California, Berkeley

Authors’ Note

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

Correspondence concerning this article should be addressed to Minah H. Jung, 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.
Abstract

People often make judgments about their own and others’ valuations and preferences. Across 12 studies (N=17,939), we find a robust bias in these judgments such that people tend to believe that others have more intense experiences than they do, leading to overestimation of others’ valuations and preferences. We argue that this overestimation arises because estimations of others’ preferences rely on people’s intuitive, core representations of the experience itself (i.e., whether the experience is positive or negative). We first demonstrate that the overestimation bias is pervasive for a wide range of positive (Studies 1-4) and negative experiences (Study 5), and is not merely an artifact of how preferences are measured (Study 6). This overestimation bias ultimately forms a paradox in how people think that others tradeoff between valuation and utility (Study 7). Specifically, people believe that an identically-paying other would enjoy the same experience more than they would, but also that an identically-enjoying other would pay more for the same experience. Such paradoxical judgments do not extend to domains unrelated to preference and valuation (Studies 8A-8B), but do extend to other preference measures, such as willingness-to-wait (Studies 9-10). 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 11). These findings suggest that social judgments of others’ preferences are not only largely biased, but they also ignore how others make tradeoffs between evaluative metrics.

Keywords: Overestimation bias, comparative judgments, valuation, preferences, paradox
“Overestimating the Valuations of Others: People Perceive Others as Experiencing Everything More Intensely”

Even the most enjoyable experiences can have shortcomings. A first trip to Paris may include a panoramic view from the top of the Eiffel Tower, a tour of the world-famous Louvre, and relaxing strolls along the Seine. 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 art, or simply dislikes strolling in general. Accordingly, although the traveler can anticipate their experience to be positive, they do 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 Paris 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 a vacation in Paris. 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 paper, 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.
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 (see Krueger, 1998 for a review on this tendency). 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). Evaluations of college graduation are, for example, somewhat complicated. On the one hand, people feel the happiness of achieving a life goal, but on the other hand they may also feel the sadness of parting ways with friends. Such an experience can therefore be anticipated to have a bivariate emotional profile. College graduations may be prominently identified as mixed, but other emotional experiences might be seen as primarily univariate.
Consider the person anticipating the Paris vacation. They are likely aware of the bivariate nature of the experience, recognizing that even if there are many positives to the trip, there will still be some negatives as well (e.g., the cost of the trip and 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”). The person traveling to Paris will tend to think that fellow travelers will have a more uniformly positive experience. Indeed, recent research suggests that one way that people manage the difficulty of predicting others’ preferences is by oversimplifying them. For example, people assume that others’ liking for one product indicates disliking for a dissimilar product (Barasz, Kim, & John, 2016).

This prediction is consistent with psychological mechanisms considering the 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). In particular, past theorizing on intuitive confidence (Simmons & Nelson, 2006) indicates that the ease with which people generate an initial judgment informs the extent to which they try to subsequently correct it. 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 extreme evaluations over the same stimuli. In this way, the Paris traveler will not only overestimate the number of people who will choose Paris over Rome as a destination, but they will believe that other people will also enjoy Paris 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., they personally stay quiet because they are embarrassed to ask something foolish, but they think other people stay quiet because they understood the lecture). Thoughts about the valuations of others may follow a similar pattern. Although people have insight into their own mixed feelings, 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, there are some related findings that are consistent with it. We note that all of these findings are consistent with our general predictions, but each is thought to be explained by largely independent mechanisms. For example, most notably, 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 eight ounces 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 two inches taller ($191 vs. $895). This
finding is suggested to arise in part from beliefs about how other people value money.\(^\text{1}\) 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. In addition, moving beyond the domain of specific products, people also seem to think that other people are getting more enjoyment out of social life, by attending more parties and events (Deri, Davidai, & Gilovich, 2017). This finding is suggested to arise from a tendency to think of atypical exemplars rather than a true average. Though all of these mechanisms (valuation of money, empathy gaps, or salient standards of comparison) certainly play a role in comparative judgments of valuation, we suggest that there may be a deeper and more general belief that underlies all of these: People simply believe that other people have more uniformly valenced experiences.

Our paper aims to document valuation overestimation, 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, practicality, and familiarity (Study 3). 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; 1)

\(^{\text{1}}\) 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 = 1943, see Study S1 in the 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. Moreover, using a continuous scale as in our studies rather than the binary choice in Frederick (2012)’s 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 supplement.
Study 4). 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 5). Moreover, this bias is not merely due to an artifact of how we measure preference (Study 6). 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. Valuation overestimation persists even in this restricted condition, leading to a distinctive error in rationality (Study 7) that is peculiar to estimates of valuation, rather than of social judgments more generally (Studies 8A-8B). We show that these effects persist across measures of valuation that have different psychological properties (Studies 9 & 10). 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 considering a broader distribution of others (both those who like and dislike the experiences), this overestimation is attenuated, and at times, eliminated (Study 11).

For all studies, we determined a stopping rule for data collection prior to data collection. For Study 2 (the field study), we collected as many participants as possible in the month of data collection. For all other studies, we determined our sample sizes before beginning data collection. We report all data exclusions, all manipulations, and all measures (Simmons, Nelson, & Simonsohn, 2013). We pre-registered Studies 1, 3-7, and 9-11. Data, R code, and pre-registrations for all studies reported in the main text and the Supplemental Materials are available at https://osf.io/b8253/.

**Analysis Plan**
Some of our studies investigated self-other differences in evaluations between-subjects (Studies 1-3, 5, & 9), 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, 4, 6-8B, & 10-11). 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 utilities. 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 spreads 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 people will think others would not only pay more for the movie, but also that others would enjoy the movie more than they would.

**Method**

**Participants and design.** Participants (N = 1653) recruited by Luth Research completed an online study for payment. Participants were randomly assigned to a 2 (person: self or other) × 2 (DV: enjoyment or WTP) between-subjects design. Only participants who answered the attention check correctly (N = 1311) were included in the analyses (pre-registration: https://osf.io/k4xj2/).

**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 (2012)’s 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. On the other hand, when estimating how much another person might enjoy the movie, participants may have 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) × 2 (enjoyment: self or other) × 2 (DV order: WTP first or enjoyment first) × 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 6pm to 9:30pm 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.  

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. 20% of your payment goes to [a local charity name and the description of the charity].” This instruction was also provided on the questionnaire.

---

2 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 structures, 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 to their own. And in fact, research has shown that people believe that they will be more generous than others in charitable contexts (Epley & Dunning, 2000). However, more recent research examining voluntary payments—in pay-what-you-want and pay-it-forward settings—provided evidence that people believe that others will pay more than they themselves do in these settings (Jung, Nelson, Gneezy, & Gneezy, 2014). Thus, Study 2 aimed to clarify whether people overestimate others’ payments even 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.
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: (1) other payment, other enjoyment, self payment, (2) other enjoyment, other payment, self payment, (3) self payment, self enjoyment, other payment, and (4) self enjoyment, self payment, other payment (see Figure S5 in the 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. 202 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 neither Person Order nor DV Order for these effects.

Our second WTP analysis compared customers self-reported payments to their perceptions of others’ payments. 236 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., $5 per ticket). Since that error is more likely to apply to self-payments than other-payments, the comparison using reported payments is somewhat conservative. Nevertheless, we found that the results from self-reported payments were consistent with the results from customers’ invoices: People thought that other customers would pay more ($M = $7.11, $SD = $3.10) than they reported paying themselves ($M = $6.21, $SD = $3.57), one-sample $t(235) = 4.45, p < .001, d = .29$. There were no interactions with either Person Order nor 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. Replicating Study 1, people ($N = 241$ complete responses) 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 relevant—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 (pre-registration: https://osf.io/6djkh/).
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 an 11-point scale (0 = not at all; 10 = 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).

Table 1. WTP for self versus other in Study 3.

<table>
<thead>
<tr>
<th>Product</th>
<th>Self (n = 187)</th>
<th>Other (n = 189)</th>
<th>t</th>
<th>Cohen’s d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treadmill</td>
<td>$279.34 ($222.90)</td>
<td>$373.98 ($255.19)</td>
<td>3.83***</td>
<td>.40</td>
</tr>
<tr>
<td>Toyota Prius</td>
<td>$19796.24 ($9035.65)</td>
<td>$22506.54 ($7668.10)</td>
<td>3.02**</td>
<td>.31</td>
</tr>
</tbody>
</table>

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.
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 utilities 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.

Table 2. Enjoyment for self versus other in Study 3.

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

*Note. Standard deviations are in parentheses. * p < .05, ** p < .01, *** p < .001. (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.)*

### Study 4: Within-Subjects

Studies 1-3 used between-subject 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 4, we test whether the observed overestimation bias still persists in a within-subject 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 = 1628) recruited by Luth Research completed an online study for payment. Participants were randomly assigned to a 2 (person: self or other) × 2 (DV: WTP or enjoyment) mixed-factors design with Person as a within-subjects factor, and DV as a between-subjects factor. Only participants who answered the attention check correctly (N = 1438) were included in the analyses (pre-registration: https://osf.io/6wz4r/).

**Procedure.** Participants evaluated a 12-piece box of Godiva chocolates. 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 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 chocolates than they themselves would pay ($M_{\text{Difference Score}} = $3.20, $SD_{\text{Difference Score}} = $5.34), one-sample $t(740) = 16.26, p < .001, d = .60$. At every level of WTP, participants estimated that others’ WTP would be higher than their own (see Figure 2A).

Moreover, within-subjects, participants also expected that others would enjoy the box of chocolates significantly more ($M = 80.43, SD = 20.37$) than they themselves would ($M = 75.86, SD = 29.09$), one-sample $t(696) = 5.57, p < .001, d = .21$. At every level of enjoyment, participants thought that others would derive more enjoyment from the box of chocolates (see Figure 2B). These results confirm that the overestimation bias appears regardless of whether it is measured in a between-subject or a within-subjects design.

(A)
Figure 2. Self and other preference for a box of chocolates in Study 4. 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 the 95th percentile. Panel B shows self and other anticipated enjoyment.

Study 5: Disliking
Studies 1-4 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 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 intense 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.

Though that is our prediction, it is certainly not the only possible prediction. 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. Because existing products tend to be considered positive (e.g., a trip to Paris) or neutral (e.g., a tube of toothpaste), Studies 1-4 cannot differentiate between these two explanations. Thus, in Study 5, 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 = 3291) recruited from MTurk completed an online study for payment. Participants were randomly assigned to a 2 (person: self or other) × 2 (DV: WTP to avoid or dislike) × 8 (negative experiences) mixed design study with negative
experiences as a within-subject factor. Only participants who answered the attention check correctly (N = 3096) were included in the analyses (pre-registration: https://osf.io/q2ukx/).

**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, 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, 100 = extremely).

We used the eight negative experiences in Appendix J of Frederick (2012). The negative experiences were: (a) eating four hardboiled eggs over 15 minutes, (b) walking around your neighborhood with a shirt that says “Repent”, (c) sharpening 100 pencils using a mechanical sharpener, (d) going without sleep tonight, (e) fast 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 $t(1532) = 4.32, p < .001, d = .22$ (see Table 3 for results by experience), and that others would dislike the negative experiences more than they themselves would, aggregate $t(1560) = 4.71, p < .001, d = .24$ (see Table 4 for results by experience).

---

4 Aggregate scores were calculated in the same way as in Study 3 (see Footnote 3).
Table 3. *WTP to avoid for self and other in Study 5.*

<table>
<thead>
<tr>
<th>Negative Experiences</th>
<th>Self</th>
<th>Other</th>
<th>t</th>
<th>Cohen’s d</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n = 758)</td>
<td>(n = 776)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sharpening 100 pencils using a mechanical sharpener</td>
<td>$7.18</td>
<td>$9.79</td>
<td>4.17***</td>
<td>.21</td>
</tr>
<tr>
<td></td>
<td>($11.84)</td>
<td>($12.71)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fast for four consecutive days</td>
<td>$96.94</td>
<td>$123.27</td>
<td>3.63***</td>
<td>.19</td>
</tr>
<tr>
<td></td>
<td>($141.54)</td>
<td>($142.39)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Go without sleep tonight</td>
<td>$33.53</td>
<td>$41.53</td>
<td>3.17**</td>
<td>.16</td>
</tr>
<tr>
<td></td>
<td>($48.21)</td>
<td>($50.78)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walk around your neighborhood with a shirt that says “Repent”</td>
<td>$18.02</td>
<td>$22.47</td>
<td>2.89**</td>
<td>.15</td>
</tr>
<tr>
<td></td>
<td>($30.18)</td>
<td>($30.23)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Composing and reciting a poem about love</td>
<td>$7.35</td>
<td>$8.98</td>
<td>2.33*</td>
<td>.12</td>
</tr>
<tr>
<td></td>
<td>($13.59)</td>
<td>($13.69)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eat hardboiled eggs over 15 minutes</td>
<td>$6.61</td>
<td>$7.84</td>
<td>1.86</td>
<td>.09</td>
</tr>
<tr>
<td></td>
<td>($13.36)</td>
<td>($12.33)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eat 3 large jalapeno peppers with no drink</td>
<td>$34.61</td>
<td>$38.69</td>
<td>1.54</td>
<td>.08</td>
</tr>
<tr>
<td></td>
<td>($54.24)</td>
<td>($49.76)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shave your head completely</td>
<td>$187.08</td>
<td>$212.31</td>
<td>1.64</td>
<td>.08</td>
</tr>
<tr>
<td></td>
<td>($304.46)</td>
<td>($298.07)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>4.32***</td>
<td>.22</td>
</tr>
</tbody>
</table>

*Note.* Standard deviations are in parentheses. * p < .05, ** p < .01, *** p < .001

Table 4. *Dislike for self and other in Study 5.*

<table>
<thead>
<tr>
<th>Negative Experiences</th>
<th>Self</th>
<th>Other</th>
<th>t</th>
<th>Cohen’s d</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n = 799)</td>
<td>(n = 763)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sharpening 100 pencils using a mechanical sharpener</td>
<td>47.53</td>
<td>55.11</td>
<td>5.09***</td>
<td>.26</td>
</tr>
<tr>
<td></td>
<td>(30.02)</td>
<td>(28.82)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fast for four consecutive days</td>
<td>78.99</td>
<td>85.37</td>
<td>4.97***</td>
<td>.25</td>
</tr>
<tr>
<td></td>
<td>(28.09)</td>
<td>(22.11)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Go without sleep tonight</td>
<td>78.61</td>
<td>77.50</td>
<td>-.83</td>
<td>-.04</td>
</tr>
<tr>
<td></td>
<td>(27.20)</td>
<td>(25.63)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Activity Description</td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>p-value</td>
<td>p-value</td>
</tr>
<tr>
<td>-------------------------------------------------------------------------------------</td>
<td>-----------</td>
<td>-----------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>Walk around your neighborhood with a shirt that says “Repent”</td>
<td>63.52 (33.68)</td>
<td>63.65 (26.20)</td>
<td>.08</td>
<td>.00</td>
</tr>
<tr>
<td>Composing and reciting a poem about love</td>
<td>52.14 (32.78)</td>
<td>53.96 (27.18)</td>
<td>1.20</td>
<td>.06</td>
</tr>
<tr>
<td>Eat hardboiled eggs over 15 minutes</td>
<td>44.33 (35.04)</td>
<td>49.23 (25.77)</td>
<td>3.14**</td>
<td>.16</td>
</tr>
<tr>
<td>Eat 3 large jalapeno peppers with no drink</td>
<td>73.28 (31.71)</td>
<td>76.72 (24.88)</td>
<td>2.38*</td>
<td>.12</td>
</tr>
<tr>
<td>Shave your head completely</td>
<td>71.67 (34.02)</td>
<td>77.40 (22.37)</td>
<td>3.92***</td>
<td>.20</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>4.71***</td>
<td>.24</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note. Standard deviations are in parentheses. *p < .05, **p < .01, ***p < .001

These results support the notion that people do not simply think that others generally
have more positive experiences than they do. Instead, for both positive and negative experiences,
people believe that others have more intense experiences than they do.

**Study 6: Different Scales**

In our studies thus far, we consistently found that people overestimated others’ liking for
positive experiences and others’ dislike 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 6, we try to correct this potential confound by using mismatched and bipolar scales. Perhaps, by allowing people to report mixed feelings towards a stimulus, they will stop seeing others as having more intense experiences.

Method

Participants and design. Participants (N = 1215) recruited from MTurk completed an online study for payment. Participants were randomly assigned to a 2 (person: self or other) × 2 (experience: positive or negative) × 3 (scale: like, dislike, or bipolar) mixed design with Person as the only within-subjects factor. Only participants who answered the attention check correctly (N = 1095) were included in the analyses (pre-registration: https://osf.io/79jsy/).

Procedure. Participants in the positive experience conditions evaluated a 30-minute 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 ($M = 76.86, SD = 28.73$), one-sample $t(146) = 2.68, p = .008, d = .22$ (see Table 5 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 measured. This seems to rule out the potential artifactual
an explanation for the overestimation bias. Even when using mismatched scales or bipolar scales, people believe that others consume positive and negative experiences more intensely than they themselves do.

Table 5. Self and other ratings of preference using different scales in Study 6.

<table>
<thead>
<tr>
<th>Product</th>
<th>Like Self</th>
<th>Other</th>
<th>One-sample t</th>
<th>Dislike Self</th>
<th>Other</th>
<th>One-sample t</th>
<th>Bipolar Self</th>
<th>Other</th>
<th>One-sample t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Massage (Positive)</td>
<td>76.86</td>
<td>81.95</td>
<td>2.68**</td>
<td>23.52</td>
<td>14.16</td>
<td>-4.97***</td>
<td>27.12</td>
<td>35.37</td>
<td>4.56***</td>
</tr>
<tr>
<td>(28.73)</td>
<td>(16.33)</td>
<td></td>
<td></td>
<td>(31.56)</td>
<td>(19.02)</td>
<td></td>
<td>(27.34)</td>
<td>(13.99)</td>
<td></td>
</tr>
<tr>
<td>Sharpening pencils (Negative)</td>
<td>28.41</td>
<td>23.53</td>
<td>-4.10***</td>
<td>56.68</td>
<td>66.48</td>
<td>6.48***</td>
<td>-12.63</td>
<td>-21.09</td>
<td>-6.34***</td>
</tr>
<tr>
<td>(27.08)</td>
<td>(22.71)</td>
<td></td>
<td></td>
<td>(30.15)</td>
<td>(26.06)</td>
<td></td>
<td>(24.44)</td>
<td>(20.84)</td>
<td></td>
</tr>
</tbody>
</table>

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

Study 7: Others with the Identical Value or Preferences

Studies 1-6 provide evidence that people see others as having more intense experiences. We argue that this overestimation is largely driven by a tendency to evaluate others’ experiences as more uniformly consistent with the core representation of the experience. Though the results are consistent with that account, an alternative account might be that the bias in not in the inference about experience as we propose, but rather in what kind of other person people are considering.

When asked to consider others, it is possible that people 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. Our next set of experiments test this explanation by clarifying the target of comparison. Specifically, we gave people a very understandable point of reference: themselves. We asked participants to consider another person who is identical on one dimension of experience (e.g., a person who is willing to pay an identical amount) and then ask how that person would respond on the alternative dimension of experience (e.g., anticipated enjoyment). If the overestimation effect derives entirely from the consideration of an exceptional prototypical other, then considering this situation should eliminate the effect.

If, on the other hand, people intuitively overweight the core representation of the experience (e.g., chocolates are good) when judging others, and apply that judgment too broadly, then clarifying the target of evaluation will not reduce the bias. Using this new experimental paradigm, we provide a stringent test of this possibility in Study 7.

Method

Participants and design. Participants (N = 1798) recruited by Luth Research completed an online study for payment. Participants were randomly assigned to a 2 (matched: WTP or
enjoyment) × 3 (product) between-subjects design. Only participants who answered the attention check correctly (N = 1121) were included in the analyses (pre-registration: https://osf.io/f42r3/).

**Procedure.** Participants in the *matched WTP* conditions were first asked how much they would be willing to pay for a product. Next, they rated how much they would like to have the product 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 that they indicated for the same product. Finally, they rated how much they thought that person would like to have the same product on the same 101-point scale.

Participants in the *matched enjoyment* conditions were first asked how much they would like to have a product 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 product. Then, they were asked to consider another participant completing the survey who also indicated that they would like to have the same product as much as they would. Finally, they indicated how much they thought that person would be willing to pay for the same product.

The three products were: (a) a box of Bic Round Stic ballpoint pens (black ink, 60/box), (b) a bottle of 2010 Alpha Omega Napa Valley Cabernet Sauvignon, and (c) a Green Mound Juniper bonsai tree.

**Results and Discussion**

Even when asked about another person who would like to have the product just as much as they did, participants believed that those others would be willing to pay significantly more for the product than they themselves would, aggregate one-sample \( t(596) = 8.49, p < .001, d = .35 \) (see Table 6 for results by product).^5

---

^5 Because product was between-subjects and person was within-subjects in this study, aggregate scores simply reflected the difference scores for each product combined in one variable. A one-sample t-test was then conducted to compare the aggregate difference score to 0. For WTP, these difference scores were winsorized prior to aggregation.
Table 6. **WTP for products at the same level of enjoyment in Study 7.**

<table>
<thead>
<tr>
<th>Product</th>
<th>n</th>
<th>Self</th>
<th>Other</th>
<th>One-sample t</th>
<th>Cohen’s d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bonsai tree</td>
<td>200</td>
<td>$16.91</td>
<td>$23.26</td>
<td>6.44***</td>
<td>.46</td>
</tr>
<tr>
<td></td>
<td></td>
<td>($21.34)</td>
<td>($28.55)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wine</td>
<td>196</td>
<td>$22.96</td>
<td>$28.14</td>
<td>5.19***</td>
<td>.37</td>
</tr>
<tr>
<td></td>
<td></td>
<td>($72.82)</td>
<td>($46.08)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pen</td>
<td>201</td>
<td>$12.11</td>
<td>$10.85</td>
<td>3.73***</td>
<td>.26</td>
</tr>
<tr>
<td></td>
<td></td>
<td>($44.64)</td>
<td>($13.36)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>597</td>
<td></td>
<td></td>
<td>8.49***</td>
<td>.35</td>
</tr>
</tbody>
</table>

*Note.* Means and standard deviations reflect raw scores. ***p < .001.

Even when asked about another person who would be willing to pay the same amount for the product, participants believed that those others would enjoy the product significantly more than they themselves would, aggregate one-sample \(t(523) = 10.43, p < .001, d = .46\) (see Table 7 for results by product).

Table 7. **Enjoyment for products at the same WTP in Study 7.**

<table>
<thead>
<tr>
<th>Product</th>
<th>n</th>
<th>Self</th>
<th>Other</th>
<th>One-sample t</th>
<th>Cohen’s d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bonsai tree</td>
<td>159</td>
<td>40.79</td>
<td>50.18</td>
<td>5.64***</td>
<td>.45</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(32.89)</td>
<td>(28.32)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wine</td>
<td>181</td>
<td>45.24</td>
<td>60.23</td>
<td>7.71***</td>
<td>.57</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(34.29)</td>
<td>(28.63)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pen</td>
<td>184</td>
<td>45.60</td>
<td>52.78</td>
<td>4.60***</td>
<td>.34</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(29.48)</td>
<td>(25.36)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>524</td>
<td></td>
<td></td>
<td>10.43***</td>
<td>.46</td>
</tr>
</tbody>
</table>

*Note.* Standard deviations are in parentheses. ***p < .001.

Even when evaluating someone identical on one measurement of liking, people thought other people would show more liking on an alternative measure. This pattern of results suggests that our effects cannot be purely explained by a tendency to think of an extreme other. Instead,
and in keeping with our predictions, this pattern is what would be predicted from research on intuitive confidence (Simmons & Nelson, 2006, 2018). People start with an automatic evaluation of the target (e.g., “people would like a bottle of wine”) 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 for 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 for an additional reason. 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 bottle of wine, the person who is willing to pay more is effectively placing a higher monetary value on each unit of enjoyment. Effectively, those two wine 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 are not rational in their perceptions. 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.

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, 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 Studies 8A and 8B, we examined whether or not judgments of other measures show similar irrationalities in different, non-preference, domains.

**Study 8A: Height and Weight**

Study 8A was our initial investigation into whether people can maintain stable perceptions of relationships between correlated measures in a non-value/preference domain. Height and weight are necessarily correlated measures (i.e., tall people tend to be heavier). People might be biased such that they believe that a person of a similar height may, on average, weigh more (or less) than they do. Similarly, they might believe that a person of similar weight may, on average, be taller (or shorter) than they are. Logically, it should not be the case that people overestimate the weight of others with the same height, while simultaneously overestimating the height of others with the same weight. However, if people are generally unable to hold stable ratios between correlated measures, we should also observe a paradoxical
relationship in judgments of others’ weight and height similar to the judgments of others’ enjoyment and willingness-to-pay.

**Method**

**Participants and design.** Participants (N = 411) recruited from MTurk completed an online study for payment. Participants were randomly assigned to a 2 (estimation: height or weight) × 2 (other: matched or non-matched) between-subjects design. We did not include an attention check for this study, and thus, include all participants in our analyses.

**Procedure.** Participants in the *non-matched* conditions indicated either their weight (in pounds) and the average weight of participants completing the survey or their height (in inches) and the average height of participants completing the survey. We included these conditions to determine whether people had a general intuition that they were both taller and heavier than others.

Participants in the *matched* conditions were asked to indicate both their weight and their height. Those matched on weight were asked to estimate the height of another participant completing the survey whose weight was the same as their own. Those matched on height were asked to estimate the weight of another participant completing the survey whose height was the same as their own.

**Results and Discussion**

In the *non-matched* conditions, people thought the average other (M = 65.55, SD = 12.73) would be about as tall as they are (M = 66.18, SD = 14.35), one-sample t(99) = 1.16, p = .249, d = .12. They estimated that an average person (M = 158.85, SD = 23.76) would weigh less than they would (M = 179.16, SD = 57.41), one-sample t(92) = 3.98, p < .001, d = .41. Therefore, in the aggregate, people did not have a general intuition of being taller or shorter than the average other, but they did generally believe that they themselves were heavier than the average other.
In the *matched* conditions, participants estimated others with the same weight to be taller ($M = 66.95$, $SD = 9.96$) than they are ($M = 65.67$, $SD = 10.43$), one-sample $t(107) = 3.37$, $p = .001$, $d = .32$. On the other hand, they thought that others with the same height would weigh less ($M = 154.94$, $SD = 34.32$) than they do ($M = 177.18$, $SD = 53.16$), one-sample $t(109) = 5.97$, $p < .001$, $d = .57$. Consistent with their beliefs about the general average other, participants seemed to believe that they weighed more than others given their height, but their estimation of others’ height and weight were logically consistent with this belief. That is, they held a stable ratio of height and weight for themselves compared to others.

**Study 8B: Altruism and Donation**

Study 8A suggests that people are capable of holding stable beliefs about correlated measures. Study 8B further tests whether people could maintain stable perceptions of relationships between correlated measures in a non-value/preference domain. This time, we used measures that more closely paralleled WTP and enjoyment: willingness-to-donate and kindness. Past research demonstrates that people tend to think that they are kinder than others (Epley & Dunning, 2000). Accordingly, we would expect that people would believe that other people are less kind and that they are less willing to donate to charity. Critically, we can then also assess whether people are able to hold a stable relationship between those two related constructs when matched with someone else on one and then asked to assess that person on the other. If self-other differences in perceptions of kindness behave like self-other difference in perceptions of valuation, then people should show an irrational instability in the ratio of the two measures. If, on the other hand, this overestimation bias is unique to preferences, people may have a more stable belief about the relationship between kindness and donation that translates to a stable ratio when using the matched paradigm.

**Method**
**Participants and design.** Participants (N = 404) recruited from MTurk completed an online study for payment. Participants were randomly assigned to one of two conditions in which they were matched with another participant either on the level of kindness or the amount of donation per year. Only participants who answered the attention check correctly (N = 379) were included in the analyses.

**Procedure.** All participants first indicated how kind or generous they were on a 101-point slider scale (0 = not at all to 100 = extremely) and their annual charitable or non-profit donation amount in an open text box. Then, participants estimated how another participant taking the survey would respond. Participants were randomly assigned to either be matched on: (a) donation amount, or (b) kindness and generosity. Participants matched on donation amount estimated the kindness and generosity of another participant who also indicated the same donation amount as they did. Participants matched on kindness and generosity estimated the annual donation amounts for another participant who also indicated the same level of kindness and generosity as they did.

**Results and Discussion**

When matched on donation amount, people thought that the other person was less kind and generous ($M = 62.02, SD = 21.83$) than they were, ($M = 64.44, SD = 21.99$), one-sample $t(178) = -2.05, p = .042, d = .15$. When matched on kindness and generosity, people thought that the other person would donate non-significantly more ($M = $615.53, $SD = $1806.53) than they would, ($M = $543.34, $SD = $1657.34), one-sample $t(185) = 1.59, p = .114, d = .12$. Together this suggests a pattern in which people consistently perceive others as being less kind-per-dollar donated then they are themselves. Unlike for perceptions of value, perceptions of that generosity ration are generally consistent across the different matching procedures, suggesting that there is something distinctive about perceptions of valuation.
Across Studies 8A-8B, people were generally imperfect at estimating others in non-value domains. Nevertheless, that perception was 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. In Studies 9 and 10, we further examine whether their overestimation bias and irrational beliefs about others extends to another valuation metric that has different psychological properties from WTP and enjoyment.

**Study 9: Willingness-to-Wait**

People have a general intuition that other people like things more than they do themselves. When they are matched with someone else who anticipates enjoying something equally, they need to answer a more difficult question: Would this other person be willing to pay more for the same amount of enjoyment? Rather than answer that difficult question, people seem to give the answer to a related question that comes to mind easily and is answered intuitively: Are other people, *in general*, willing to pay more than me? (Kahneman & Frederick, 2002; Simmons & Nelson, 2006; 2018). That tendency leads to a seeming paradox in estimates of other people’s WTP for enjoyment. Specifically, depending on how the question is asked, people think that others are willing to pay more or less for the same level of enjoyment.

Is this bias specific to WTP and enjoyment, or does it manifest as a general intuition of others’ preferences? In Studies 9 and 10, we examine whether this overestimation bias and paradox also emerges with a different but commonly used valuation metric: willingness-to-wait.

**Method**

**Participants and design.** Participants (N = 1228) recruited from MTurk completed an online study for payment. Participants were randomly assigned to a 2 (person: self or other) × 4
(products) between-subjects design. Only participants who answered the attention check correctly (N = 953) were included in the analyses (pre-registration: https://osf.io/dpz34/).

**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 8 for results by product and Figure 3 for a density plot of the early movie experience). These results suggest that the overestimation bias appears more broadly in the estimation of others’ values and preferences.

Table 8. *Willingness-to-wait in minutes for products and experiences in Study 8.*

<table>
<thead>
<tr>
<th>Products</th>
<th>Self</th>
<th>Other</th>
<th>$t$</th>
<th>Cohen’s $d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early screening of the newest Star</td>
<td>42.40</td>
<td>99.72</td>
<td>6.52***</td>
<td>.83</td>
</tr>
<tr>
<td>Wars movie</td>
<td>(40.32)</td>
<td>(87.53)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>n = 119</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^6$ 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.
An autographed jersey of Michael Jordan  
98.77 (110.21) 160.93 (136.84) 3.92*** .50
n = 123 n = 122

A newly released virtual reality headset  
47.90 (70.63) 89.53 (94.91) 3.79*** .50
n = 116 n = 116

A tour of the Burj Khalifa skyscraper in Dubai  
60.73 (39.71) 73.21 (47.95) 2.15* .28
n = 117 n = 113

Total  
8.14*** .53

Note. Standard deviations are in parentheses. * p < .05, ** p < .01, *** p < .001

Figure 3. Density plots for self and other wait times for an early movie screening in Study 10. Means and error bars are shown above each graph.

Study 10: Match paradigm with waiting
People overestimate others’ valuation and utility using three different valuation metrics: WTP, liking, and willingness-to-wait. In Study 10, we test whether this overestimation bias about others again leads to paradoxes about others’ valuation and utilities measured across those three metrics. In other words, we intended to investigate whether using the match paradigm in Study 7 with another measure of preference (wait time) also leads to paradoxical beliefs about others’ preferences.\(^7\)

**Method**

**Participants and design.** Participants (N = 3267) recruited from MTurk completed an online study for payment. Participants were randomly assigned to one of four between-subject conditions: (1) enjoyment at matched wait time (n = 617), (2) WTP at matched wait time (n = 647), (3) wait time at matched enjoyment (n = 791), and (4) 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 = 2782) were included in the analyses (pre-registration: https://osf.io/fqah7/).

**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).

\(^7\) We initially conducted a study (N=1188) examining this question without pre-registering and using a smaller sample (see Study S2 in the Supplemental Studies and Materials). We conducted Study 10 to confirm the results of this study with a larger sample.
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 7, we demonstrated the value tradeoff 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 did, 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 values when people are making judgements about the valuation and preferences of others.

**Study 11: 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 people’s estimation of others’ valuations and preferences rely on their core representations of given products and experiences. One potential way to disrupt this 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 Paris vacationer brings to mind all the positive features of the city, but thinking of a less-positive Paris vacationer might remind people of exactly those less salient features which are not necessarily so positive. Study 11 tested whether asking people to estimate a distribution of others’ preferences, rather than a point estimate of others’ preferences, attenuates the overestimation bias.

**Method**

**Participants and design.** Participants (\(N = 1968\)) 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 (measure: WTP or liking) \(\times\) 2 (format: point estimate or distribution) mixed-design study with Person as the only within-subjects factor. Only participants who answered the attention check correctly (\(N = 1709\)) were included in the analyses (pre-registration: https://osf.io/6238v/).
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 (WTP condition) or their liking on a scale that ranged from 0 (not at all) to 10 (extremely). 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 condition) or estimated others’ liking on the same 0 to 10 scale as self evaluations. In the distribution conditions, participants were asked to allocate the percentage of the other survey takers who would choose each of the 11 different choice options, totaling 100%. 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 condition) 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 condition).

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) × 2 (order: self first or other first) × 2 (format: point estimate or distribution) mixed-model ANOVA, with Person as the only within-factor.
We observed a significant main effect of Person, $F(1, 885) = 74.31, p < .001, \eta^2_p = .08$, 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. Consistent with this possibility, the Person $\times$ Format interaction was significant, $F(1, 885) = 42.18, p < .001, \eta^2_p = .05$. 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 $t(481) = 11.00, p < .001, d = .50$ (see Figure 4). 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 5). This attenuation did not significantly change based on which person they evaluated first (self or other), $F(1, 885) = 2.11, p = .147, \eta^2_p = .002$.

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 point estimate condition in Study.
11. 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 point estimate condition. Similarly, the estimated percentage was calculated as the proportion of participants who guessed that the average other would pay the specific price out of all the participants in the WTP point estimate condition.

![WTP Distribution](image)

*Figure 5.* 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 11. 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).

**Liking.** As with the 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) × 2 (order: self first or other first) × 2 (format: point estimate or distribution) mixed-model ANOVA, with Person as the only within-factor. We observed a significant main effect of Person, $F(1, 812) = 250.63, p < .001, \eta^2_p = .24$, 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 × Format interaction was significant, \(F(1, 812) = 10.06, p = .002, \eta^2_p = .012\). 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 \(t(406) = 14.20, p < .001, d = .70\) (see Figure 6). When estimating others with a distribution, this effect was 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 7). 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), \(F(1, 812) = .97, p = .326, \eta^2_p = .001\).

**Figure 6.** Actual and estimated percentage of participants who would like a card that buys two drinks at Starbucks a given amount in the liking point estimate condition in Study 11. Actual percentage was calculated as the proportion of participants who indicated that they would like the card that specific amount out of all the participants in the liking point estimate condition. Similarly, the estimated percentage was calculated as the proportion of participants who guessed that the average other would like the specific amount out of all the participants in the liking point estimate condition.
Figure 7. 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 11. 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).

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.

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 value goods more than they do (Studies 1-6, & 9). 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 Paris is generally thought to be positive, 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., Paris 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 (Study 7). 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). This paradoxical belief is unique to preferences (Study 8A-8B), but extends to measures of preference beyond WTP and enjoyment (i.e., willingness-to-wait in line; Studies 10). 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 11).

**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 versus 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 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 7 and 10 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, 1979; 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 (Daniel 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 for the overestimation bias**

This paper reports 12 experiments showing 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 the expression of 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 Supplemental Materials.

**Others with Extreme Preferences.** Study 7 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 Supplemental Materials.
First, in Study S3 (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 US person’s evaluations of a Star Wars product. Though these Star Wars fans rationally understood that the average US 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 Study S4 (N = 1214), we used the same paradigm as in Study 7 with an additional factor (Other). In addition to examining how people view identical others (i.e., those matched on either enjoyment or WTP), we explored people’s estimations 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 5 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 5 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 7 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 5 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 Study 7 that the bias cannot be fully explained by the salience of others with extreme preferences or the extremity of people’s own preferences. When people explicitly consider others who are less positive towards a product, people display rational responses. However, when considering average others or those who should have similar preferences, the overestimation bias persists.

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 co-worker 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 here 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 tradeoff 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.
References


OVERESTIMATION BIAS IN OTHERS’ VALUATION

https://doi.org/10.1146/annurev.psych.43.1.87

https://doi.org/10.1016/0022-1031(77)90049-X

https://doi.org/10.1037/0096-3445.135.3.409


https://doi.org/10.2139/ssrn.2205186


https://doi.org/10.1037/0003-066X.50.5.364


https://doi.org/10.2307/2937956