There's No I in Expectations:

The Effect of Expectations as Reference Points on

the Recommendation Likelihood of Experiential Purchases

STEPHANIE TULLY

AMAR CHEEMA

ON AMIR

DAVIDE PROSERPIO

This Draft: September 2018

Please do not cite or quote without prior permission

Author Note: Stephanie M. Tully (corresponding author: phone: (213) 821-4209, email: stephanie.tully@marshall.usc.edu) and Davide Proserpio (email: proserpi@marshall.usc.edu) are Assistant Professors of Marketing, Marshall School of Business, University of Southern California. Amar Cheema (email: cheema@virginia.edu) is the O'Dell Professor of Commerce at the McIntire School of Commerce, University of Virginia. On Amir (email: oamir@ucsd.edu) is Professor of Marketing, Rady School of Management, University of California San Diego.

There's No I in Expectations:

The Effect of Expectations as Reference Points on

the Recommendation Likelihood of Experiential Purchases

Contribution Statement

Experiential purchases are a large and growing part of the economy. Existing literature on experiences has examined how different components and structures of an experience influence retrospective evaluations, and how experiences are evaluated relative to material goods. However, less is known about how factors that occur after an experience is over may impact evaluations. The primary objective of the current research is to understand whether and how, holding one's experience constant, the use of different types of reference points can systematically influence consumers' evaluations of experiential purchases. Specifically, we examine whether and why encouraging the use of expectations as a reference point changes willingness to recommend an experiential purchase to others.

Despite existing research suggesting that expectations are largely the dominant reference point used by consumers in forming evaluations, we suggest that consumers do not readily consider their expectations when deciding whether to recommend an experiential purchase. Moreover, this research identifies two distinct components of experiential purchases: what is offered to the consumer and what is experienced by the consumer. We argue that expectations shift the focus consumers place on these two components, which in turn systematically decreases willingness to recommend experiential purchases. We differentiate our proposed process from competing explanations. For example, the reduction in recommendations is not because expectations are simply high reference points (and thus make everything worse) as the effect is unique to experiential purchases and does not extend to material goods.

In addition to contributing to the literature on evaluations of experiences and on the literature about material goods versus experiences, this work also contributes to research on reference points and framing effects. Much of the existing research on reference points has focused on differences related to the height or strength of one's reference point. In contrast, this research is among the first to consider how horizontally differentiated reference points may systematically impact decision making.

There's No I in Expectations:

The Effect of Expectations as Reference Points on

the Recommendation Likelihood of Experiential Purchases

Abstract

This research investigates whether and how reference points systematically change consumers' willingness to recommend experiential purchases to others. Across lab studies, field studies, and millions of online reviews, we demonstrate that when reflecting on experiential purchases, considering one's expectations systematically decreases individuals' likelihood of recommending such purchases to others. This happens when consumers consider expectations after an experience occurs and is not the result of changes to the experience itself. Instead, we suggest that consumers typically rely on their subjective experience (e.g., emotional reactions) when determining whether to recommend an experience to others. However, when used as the reference point, expectations shift consumers' focus away from their subjective experience toward the external attributes of the experience (e.g., the setting, what is consumed), which reduces willingness to recommend the experience. In line with this proposed explanation, we demonstrate that this effect is unique to experiential purchases and does not extend to material purchases, which are less likely to be evaluated based on one's subjective experience to begin with. Our research suggests that when considering whether to recommend experiential purchases, expectations are not the default reference point. We discuss this conclusion and other implications of this work.

Keywords: evaluations of experiences, hedonic experiences, expectations, reference points, satisfaction, experiential purchases

Americans spend an incredible amount of discretionary money on experiential purchases. In 2013, they spent $22 billion dollars on bars and nightclubs, $140 billion on hotels, bed and breakfasts, and hostels, $15 billion at amusement parks, and well over $400 billion eating at restaurants (IBISWorld Inc., 2013). Word of mouth is thought to be of particular importance to the success of companies offering experiential purchases since an experience's intangible nature makes it difficult to evaluate quality in advance of purchase (Litvin, Goldsmith, and Pan 2008). Indeed, research has shown that word of mouth impacts future sales across a range of experiential purchases (Liu, 2006; Luca, 2016; Neilsen, 2014; Rui, Liu, and Whinston 2013). As such, companies often go to great lengths to encourage consumers to engage in word of mouth (Kumar, Peterson, and Leone 2007). In doing so, these companies may knowingly or inadvertently cue consumers to different reference points. For instance, companies requesting reviews may ask, "Did this purchase meet your expectations? Tell others about it!" While it is possible that willingness to recommend a purchase may differ simply by encouraging people to use different types of reference points, whether this is true, and why such differences may occur is currently an open question.

In this research, we consider how decisions to recommend experiential purchases are formed and examine whether and how the use of different reference points impacts subsequent recommendations. Specifically, we investigate whether encouraging consumers to consider their expectations for an experiential purchase predictably impacts recommendation likelihood. Despite existing research suggesting that expectations are largely the dominant reference point used by consumers in forming evaluations (e.g., Oliver 1980), we suggest that consumers do not readily consider their expectations in deciding whether to recommend an experiential purchase to others. Moreover, we identify two distinct components of experiential purchases: what is offered

to the consumer and what is experienced by the consumer. We argue that the consideration of expectations shifts the focus consumers place on these two components, which in turn systematically decreases willingness to recommend experiential purchases.

We demonstrate our effect and find support for our proposed process across a variety of lab and field studies. The consideration of expectations makes consumers less likely to recommend the restaurants where they have eaten, college courses they have taken, and movies they have seen. This effect is not isolated to specific experiential domains as it replicates when examining idiosyncratic consumer-generated experiential purchases. Furthermore, the reduction in recommendations is not because expectations are simply high reference points and thus make everything worse as the effect is unique to experiential purchases and does not extend to material goods. Instead, we provide evidence that expectations change the way experiential purchases are conceptualized by shifting focus away from one's personal experience and emotional reactions. Finally, using millions of real consumer reviews from two popular travel related content websites, we replicate our proposed effects in archival data. We show on an aggregate level that the consideration of expectations reduces both consumers' individual star-ratings and the proportion of consumers who will recommend a hotel to others.

THEORETICAL BACKGROUND

A large body of research on evaluations of experiences considers how different components and compositions of an experience affect subsequent evaluations (e.g. Ariely 1998; Kahneman et al. 1993; Redelmeier and Kahneman 1996). In the current research, we do not examine changes to the experience itself, but rather examine how the use of different reference

points at the time of evaluation impacts willingness to recommend experiential purchases. Prior work reveals that the evaluation of a stimulus depends not only on characteristics of the stimulus but also on the reference point from which it is evaluated (e.g., Kahneman and Tversky 1979). While much work on reference points examines how differences in the magnitude or strength of a reference point impact decision making, reference points can differ in type as well. Indeed, consumers may have multiple reference points to consider for any single purchase (Kahneman 1992; Ordonez, Connolly, and Coughlan 2000). In addition to variation in which reference point consumers choose to use, reference points can be externally imposed or encouraged. Indeed, in word-of-mouth promotion requests, companies can cue consumers to specific reference points (e.g., "Did we meet your expectations?"). Although research has recognized that multiple types of reference points can be used by consumers, it is seemingly unknown whether and how the use of certain reference points impacts consumers' evaluations in any systematic way. This is the focus of the current investigation.

In this research, we examine how the use of expectations as reference points impacts consumers' willingness to recommend experiential purchases. The majority of past research suggests that eliciting expectations *after* an experience is over should not affect evaluations. Expectations are believed to commonly serve as frames of reference against which experiences are regularly judged (e.g., Anderson 1973; Deighton 1984; Hoch and Ha 1986; Oliver 1980). Indeed, prior research argues that consumers inherently judge experiences against expectations to form satisfaction judgments (e.g., Johnson, Anderson, and Fornell 1995). Moreover, expectations measured post-consumption are believed to be more influential in satisfaction judgments than pre-consumption expectations (e.g., Zwick, Pieters, and Baumgartner 1995). Thus, if consumers

naturally use their expectations as reference points, encouraging the consideration of expectations should have no effect.

More recently, Ofir and Simonson (2001, 2007) have shown that stating expectations prior to an experience significantly decreases evaluations. The authors attribute these lower evaluations (from people who are asked to state expectations prior to an experience) to greater attention on the negative aspects of an experience as consumers go through the experience. That is, as people live through an experience, having their expectations salient makes people more attuned to negative deviations from their expectations leading to lower evaluations. Importantly, Ofir and Simonson suggest that the difference in satisfaction between people who state their expectations prior to an experience and people who do not do so is because retrospective memory-based evaluations are not prey to the negativity bias to the same extent as evaluations formed during an experience. Thus, Ofir and Simonson's explanation for why eliciting expectations reduces satisfaction relies on expectations being salient and accessible during an individual's actual experience. Consequently, their results suggest that encouraging the consideration of expectations *after* an experience, at the point of evaluation, should have no effect. However, in the current research, we suggest and demonstrate that encouraging the use of expectations, even after an experience is over, can have a significant negative impact on experiential purchase recommendations.

While experiential purchases are described as something purchased with the primary intention of acquiring a life experience (e.g., Van Boven and Gilovich 2003), we identify two distinct components that are integral to every experiential purchase. Specifically, we recognize a distinction between the attributes offered to a consumer and what is experienced by the consumer. For example, in the case of a restaurant experience, the experience's attributes consist

of things such as the restaurant's décor, the ambiance, the service, and the food. In contrast, what is experienced by the consumer may consist of feelings of enjoyment, satiation, or social bonding. Undoubtedly, these two components can affect one another. For example, personal taste may dictate which item a person orders on a menu or how much interaction they have with their server. Conversely, the food that is served and the ambiance of the restaurant may impact one's enjoyment. However, we suggest that the extent to which consumers focus on one component versus the other can influence their likelihood of recommending an experience.

Because the primary intention of an experiential purchase is to live through an event, when considering whether to recommend an experiential purchase, we contend that consumers primarily rely on episodic memory, which allows them to relive their experiences, including their emotional responses (Tulving 2002). Supporting this contention, previous research shows that when consumers are encouraged to think about autobiographical experiences for products (i.e., think of products as experiential purchases), participants think more about past personal experiences, show higher levels of net positive affect, and think less about specific features and attributes (Baumgartner, Sujan, and Bettman 1992; Sujan, Bettman, and Baumgartner 1993). Thus, we suggest that consumers typically consider what they personally experienced (i.e., looking inward) more so than the experience's attributes (i.e., looking outward) when determining their likelihood of recommending an experiential purchase to others.

Importantly, we suggest that expectations are more typically a function of an experience's attributes. For instance, in the case of a restaurant, one's expectations might include what will be available on the menu, what the service or ambiance will be like, or how good the food will be. In the case of a hotel, this might include the size of the room, the quality of the mattress, or the proximity to attractions. That is, we suggest that expectations are typically

conceptualized as a function of the external attributes offered to the consumer rather than the personal experience that they live through. Said differently, expectations are more likely to reflect what "it" will be like rather than what "I" will experience. If people instinctively rely on their personal experience and emotional responses when considering experiential purchases, and if expectations are typically more focused on the experience attributes, considering their expectations should shift consumers' focus from their personal experiences and emotional reactions toward the experience's attributes. How will this shift in focus impact willingness to recommend experiences?

Previous research has worked to identify factors that drive word of mouth (see Berger 2014 for a review). This research finds that people are more likely to engage in word of mouth when the content is highly emotional, arousing, and relevant to one's self-concept (e.g., Berger and Milkman, 2012; Chung and Darke 2006; Peters, Kashima, and Clark, 2009; Wojnicki and Godes 2017). Moreover, one of the motivations to engage in word of mouth is the ability to rehearse one's past experiences because doing so allows the consumer to relive their past experience and can enhance positive affect (Berger 2014; Langston 1994; Walker et. al 2009). Taken together, this research suggests that people should be more motivated to engage in word of mouth when thinking about what they personally experienced (which is more focused on the self and is more emotional in content) than they are when thinking about what the experience offered (which should be less self-relevant and lower in emotional content). Thus, to the extent expectations draw attention away from one's personal experience toward the attributes of an experience, expectations should reduce the likelihood of recommending experiential purchases to others.

In summary, we suggest that consumers typically rely more on what they personally experience (e.g., their emotional reactions) rather than the experience's attributes when deciding whether to recommend an experiential purchase to others. However, when expectations are evoked at the time of evaluation, it shifts consumers' focus away from their personal experience toward the more external attributes of the experiential purchase, which reduces recommendations of experiential purchases. More formally:

H1: The use of expectations as a reference point reduces consumers' likelihood of recommending their experiential purchases.

H2: The reduction in the likelihood of recommending experiential purchases resulting from considering one's expectations is explained by a decreased focus on one's personal experience and emotional reactions.

We differentiate our proposed process from the alternative possibility that expectations are simply unrealistically high reference points and thus decrease evaluations of anything. Specifically, our explanation relies on consumers' initial focus being on their own personal experience and emotional reactions. When this is not the case, expectations should not decrease willingness to recommend a purchase. Material purchases, which are purchases made with the primary intention of owning a tangible object (Van Boven and Gilovich 2003), are traditionally described by their attributes and product features as opposed to the experiences one intends to have with them (Baumgartner, Sujan, and Bettman 1993; Carter and Gilovich 2010). Thus, we should not expect expectations to decrease willingness to recommend material purchases. Hence:

H3: The decrease in willingness to recommend one's purchase resulting from the consideration of expectations is unique to experiential purchases and does not extend to material purchases.

We test these hypotheses in a field study, three lab studies, and two sets of archival data. In study 1, we demonstrate that restaurant patrons are less likely to recommend the restaurant they visited when encouraged to consider their expectations (versus no specific reference point or their goals) for the restaurant visit. Moreover, we show that the effect is explained by a reduced focus on one's personal experience. In study 2, we examine the generalizability of our effect by examining a different type of experiential purchase: college courses that one has taken. Even for this expensive and important experiential purchase, expectations reduce recommendation likelihood. In study 3, we provide greater evidence for the proposed process by asking consumers to draw their expectations for, goals for, or experience at a recent restaurant visit. In line with our hypothesis that expectations shift consumers' focus from their personal experience toward more external attributes of the experience, this study finds that drawings of a restaurant visit are less likely to include pictures of people (e.g., themselves) when thought of in terms of consumers' expectations. In study 4, we rule out the possibility that expectations simply reduce one's interest in recommending anything and demonstrate that expectations reduce recommendations for experiential purchases but not for material goods. Finally, we provide greater external validity by examining consumers' hotel reviews on TripAdvisor and Expedia. Using large datasets of real consumers' reviews, we find that consumers using their expectations as a reference point provide lower star-ratings and are less likely to recommend the hotel to others. These effects are explained by a reduced focus on one's emotional reactions rather than other potential explanations such as differences in hotel quality or type of reviewer.

STUDY 1: RESTAURANT RECOMMENDATIONS

In study 1, we use a field study to test whether encouraging the use of expectations as a reference point in evaluating a restaurant visit reduces willingness to recommend the restaurant to others. We conducted this study during "Restaurant Week," a week during which restaurants offer discounted 3-course meals to entice consumers to try new restaurants. Since restaurants typically make less money on each patron, the benefit of restaurant week for many restaurants is the positive word of mouth it can generate (Tuttle 2013). Thus, consumers' willingness to recommend the restaurant is critical during this time.

Moreover, we designed this study to examine and isolate any impact of using expectations as the reference point from the impact of using any positive reference point. To do so, we used a three-condition design in which we asked participants to consider their expectations for the restaurant visit, their goals for the restaurant visit, or nothing at all before indicating their willingness to recommend the restaurant to others. We expected that consumers led to consider their expectations for their restaurant visit would be less likely to recommend the restaurant relative to consumers in the other two conditions.

Method

This survey was run in conjunction with another survey being distributed for unrelated purposes, and was conducted at a number of restaurants across a town partaking in "Restaurant Week." The purveyors of the original survey had reached out to a set of 36 restaurants who agreed to give diners a card at the end of their meal. Instead of turning the survey into the server

directly, the card directed diners to complete an online survey. To incentivize the restaurant and the diners to complete the survey, each day's responses were entered into a lottery. The winner received a $50 gift card, and a similar $50 gift card was won by their server. The sample size was determined by researchers conducting the other survey. Two hundred and six restaurant patrons completed the survey. Five participants had missing data on the primary DV, leaving a final sample of 201 participants (M_{age} = 35.6 years, SD = 14.6; 62% women).

The first part of the survey was designed by the other researchers coordinating the survey collection. Thus, participants answered questions about the logistics of their restaurant visit (e.g., restaurant name, number of people they ate with, whether they had been to that restaurant before) as well as perceptions of quality and their willingness to recommend the experience (see web appendix for full list). Although administered for the purposes of the other research, these baseline ratings provided a unique opportunity to isolate the causal effect of considering different reference points. The next set of questions asked about perceptions of authenticity of the restaurant, which was the primary focus of the other researcher's survey and they are therefore not discussed here.

The questions for the current research came at the end of this survey. Following the predetermined set of questions, we randomly assigned participants to one of three reference point conditions: control, expectations, or goals. In the expectations [goals] condition, participants were asked what their expectations [goals] for the restaurant visit were (open-ended). Then, the primary dependent measure asked, "Considering your expectations [goals], how likely would you be to recommend the restaurant to a friend?" (1 = not at all likely to recommend, 7 = very likely to recommend). Additionally, participants were asked, "Considering your expectations [goals], how satisfied were you with the restaurant visit?" (1 = not at all satisfied, 7 = very

satisfied). Participants in the control condition did not receive an open-ended question and were not asked to consider any reference point while providing their likelihood of recommending the restaurant or their satisfaction. Participants then completed additional open-ended responses for the restaurant and concluded with demographics.

Results and Discussion

Recommendation Likelihood. Participants answered detailed questions at the beginning of the survey about their restaurant visit and their satisfaction. Although these ratings may make it more difficult to change evaluations (e.g., due to anchoring), they also allow us to more clearly isolate the impact of using different reference points.

We first analyzed recommendation likelihood across the three conditions (control, expectations, and goals) while adjusting for initial restaurant quality ratings (average of 10 items, see web appendix). Indeed, an ANOVA demonstrated that average rating was a significant predictor of recommendations, $F(1, 196) = 315.48$, $p < .001$. Importantly, adjusting for the differences in experience quality, there was a marginally significant effect of condition $F(2, 196) = 2.52$, $p = .083$. Planned contrasts found support for our predictions. Participants asked to consider their expectations for the experience ($M = 6.18^{1}$, SD = 0.79) were less likely to recommend the restaurant to others compared with those in the goals condition ($M = 6.44$, SD = 0.80) and those in the control condition ($M = 6.44$, SD = 0.80), $F(1, 196) = 5.04$, $p = .026$. There was no difference between those in the goals condition and those in the control, $F < 1$, NS.

---

[1] Means reported are adjusted means. Raw means are as follows: $M_{expectations} = 6.20$, SD = 1.29, $M_{goals} = 6.52$, SD = 1.04, $M_{control} = 6.32$, SD = 1.46.

Participants had also reported a recommendation likelihood at the beginning of the survey (three items, see web appendix). Thus, as an alternative analysis, we computed a difference score between the final recommendation and the initial recommendation likelihood (final- initial) to represent changes in participants' evaluation. A positive [negative] difference score thus indicates an increase [decrease] in recommendation likelihood. An ANOVA indicated a marginally significant effect of condition, $F(2, 198) = 2.55$, $p = .081$. Planned contrasts again found support for our predictions. Participants in the expectations condition ($M = -0.13$, SD = 0.68) decreased their recommendation likelihood to a greater extent than participants in the goals condition ($M = 0.24$, SD = 1.20) and those in the control condition ($M = 0.18$, SD = 1.09), $F(1, 197) = 4.89$, $p = .028$. Again, there was no difference between those in the goals condition and those in the control condition, $F < 1$, NS. See figure 1.
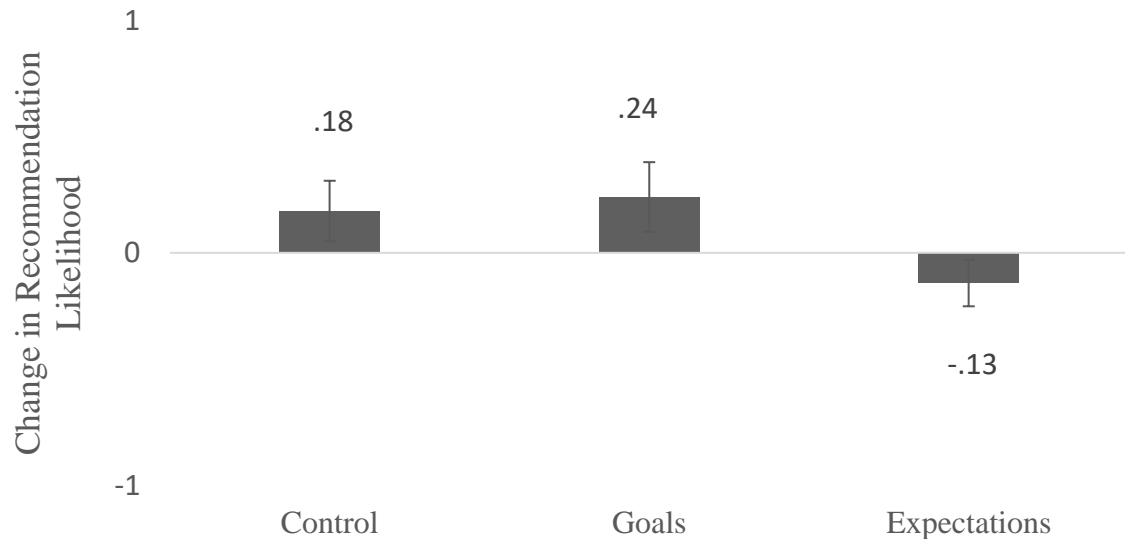
Note: Error bars represent standard error of the mean.

FIGURE 1. Change in Restaurant Recommendation Likelihood by Reference Points (Study 1).

Satisfaction. Although directionally consistent, this measure did not show any significant differences as a function of condition (adjusted means: $M_{expectations}$ = 6.24, SD = 0.81, M_{goals} = 6.30, SD = 0.81, $M_{control}$ = 6.35, SD = 0.81), all $F < 1$.

Focus on external factors versus personal experience. Only participants in the goals and expectations conditions provided open-ended responses, allowing us to examine what they were considering. Within those two conditions, not everyone completed the open-ended response. However, the 108 available responses were coded by a research assistant blind to the hypothesis for whether the written response was more focused on the restaurant's attributes (e.g., food, service, ambiance) or more focused on the participant's experience (e.g., emotional reactions, social bonding). Statements more focused on personal experience were coded as a 1 and statements more focused on the restaurant's attributes were coded as a -1. As expected, statements in the expectations condition (28.4%) were less likely to be focused on the personal experience than statements in the goals condition (78.0%), $\chi^2(1)$= 25.20, $p < .001$. Moreover, this measure of focus on one's personal experience positively predicted changes in the likelihood of recommending the restaurant from initial to final evaluations, $\beta = .328$, $t(106) = 10.14$, $p = .002$. When we regressed the change in likelihood of recommending the restaurant on reference point condition (-1 = goals, 1 = expectations), including the focus of the content (-1 = experience attributes, 1 = personal experience) significantly decreased the beta weight of the reference point condition from $\beta = -.236$, $t(106) = -2.50$, $p = .014$, to $\beta = -.122$, $t(105) = -1.15$, $p = .253$, Sobel test = 2.09, $p = .037$ in support of mediation.

Discussion. Study 1 provides initial evidence that using expectations as a reference point decreases recommendation likelihood in a field experiment. Moreover, this study suggests that the decrease in evaluations is specific to considering expectations as opposed to considering any

positive reference point. Participants asked to consider their expectations for the restaurant visit were significantly less likely to recommend the restaurant compared to those asked to consider another positive reference point (goals) or to consider nothing at all. This effect emerged even after participants were previously asked to provide ratings of their experience (which could have anchored their responses). Further, study 1 begins to provide evidence for the proposed process as the decrease in likelihood of recommending the restaurant by those considering expectations was explained by a shift in focus from one's personal experience to the experience's attributes.

A conceptual replication of this study in another restaurant field study setting using a different control condition (considering one's experience) is available in supplemental study 1 of the web appendix.

STUDY 2: COURSE RECOMMENDATIONS

In the next study, we aim to replicate our effect with a different type of experiential purchase: college courses. Thus, in this study we asked participants to consider either their goals or expectations for a course they had taken at their university before asking about their willingness to recommend the course to other students. To further rule out the possibility that decreased evaluations result from considering a "high bar," in addition to varying the reference point, we varied whether participants considered a required or elective course as we assume that people have higher expectations for elective versus required courses. Thus, if expectations decrease consumers' evaluations because expectations serve as a high reference point, evaluations of electives should suffer more than required courses. However, if expectations decrease consumers' evaluations because they shift the focus away from one's personal

experience, as we suggest they do, the reduction in evaluations should not depend on course type.

Method

This study was a 2 (reference point: goal versus expectation) x 2 (class type: required versus elective) between-subjects design. Three hundred and seventy-nine undergraduate students, as determined by the number of available lab participants, completed the computer-based study in exchange for financial compensation (M_{age}= 22.5, SD = 2.42; 66.5% male).

We asked participants to think about the first required or elective course that they had taken at the university with the type of class (elective, required) varying by condition. Next, participants in the goals condition stated their goals for the recalled course while those in the expectations condition stated their expectations. Our primary dependent measure was the likelihood of recommending the course to a friend (7-point scale: 1 = definitely not recommend, 4 = may or may not recommend, 7 = definitely recommend). After indicating their willingness to recommend the course, participants indicated how satisfied they were with the course (1 = not at all satisfied, 4 = somewhat satisfied, 7 = very satisfied).

To ensure that this manipulation indeed varied the reference points participants used to evaluate their experiential purchase, we asked participants to indicate to what extent the course met their expectations [goals] (1 = definitely did not meet my expectations [goals], 7 = definitely did meet my expectations [goals]). Finally, participants provided some information about themselves including their class year, GPA, age, gender, and native language.

Results and Discussion

Likelihood of Recommending. In line with predictions, there was a main effect of reference point on likelihood of recommending. Participants considering their expectations for the course ($M = 5.11$, SD = 1.67) were less likely to recommend it to others compared to those considering their goals for the course ($M = 5.57$, SD = 1.52), $F(1, 375) = 7.84$, $p = .005$. There was also a marginally significant main effect of class type, $F(1, 375) = 2.88$, $p = .090$. Participants were marginally more likely to recommend their course when considering an elective ($M = 5.48$, SD = 1.55) versus a required course ($M = 5.20$, SD = 1.66). However, contrary to the possibility that higher reference points result in systematically lower evaluations, the effect of reference point did not interact with course type, $F < 1$. See figure 2.
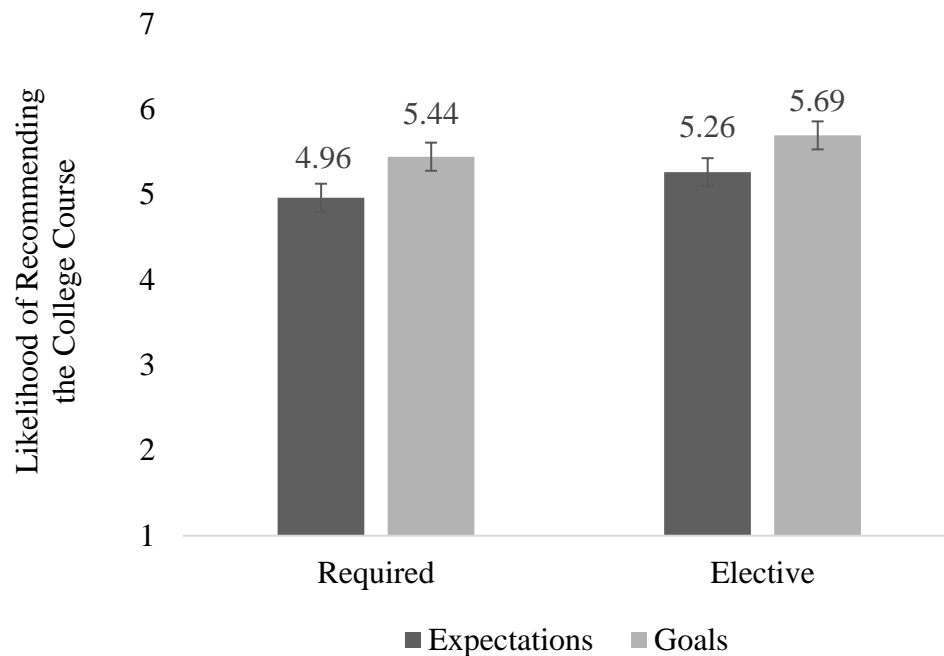
Note: Error bars represent standard error of the mean.

FIGURE 2. Likelihood of College Course Recommendation by Reference Point and Course Type (Study 2).

Satisfaction. The satisfaction measure mirrored the results of the recommendation measure. Participants considering their expectations for the course ($M = 5.07$, SD = 1.69) were less satisfied with the course compared to those considering their goals for the course ($M = 5.55$, SD = 1.30), $F(1, 375) = 9.81$, $p = .002$. Participants were more satisfied with elective courses ($M = 5.48$, SD = 1.43) compared to required courses ($M = 5.14$, SD = 1.61), $F(1, 375) = 4.74$, $p = .030$. Again, the effect of reference point did not interact with course type, $F < 1$.

Meeting the Reference Point. Substantiating our assertion that evoking goals and expectations increases their use as reference points, the results of the question asking about whether the course met their expectations [goals] mirrored those of the dependent measures. Participants indicated that the course was significantly less likely to have met expectations ($M = 5.13$, SD = 1.63) than to have met goals ($M = 5.56$, SD = 1.33), $F(1, 375) = 8.20$, $p = .004$. This measure was also marginally higher for elective courses ($M = 5.47$, SD = 1.47) than for required courses ($M = 5.21$, SD = 1.52), $F(1, 375) = 2.88$, $p = .091$. However, class type did not interact with the reference point, $F(1, 375) = 1.04$, NS.

Focus on One's Personal Experience. To examine our proposed explanation, we coded descriptions of participants' goals and expectations. Specifically, we asked a research assistant blind to the hypothesis to indicate whether the statement pertained to students' own personal experiences as operationalized through the presence of learning (e.g., understanding the topic, learning as much as possible) or performance (e.g., getting a good grade, passing the class) assertions. If the response included any such content, the research assistant coded it as a 1 and -1 otherwise. Only 46% of the open-ended responses in the expectations condition included content

about one's personal experience whereas 79% of responses in the goals condition did, $\chi^2(1)=$ 45.81, $p < .001$. Indeed, the expectations condition negatively predicted focus on one's personal experience $\beta = -.348$, $t(377) = -7.20$, $p < .001$, and focus on one's personal experience positively predicted the likelihood of recommending the course, $\beta = .207$, $t(377) = 4.12$, $p < .001$. When likelihood of recommending the course was regressed on reference point condition (-1 = goals, 1 = expectations) and course type (-1 = required, 1 = elective), including the focus on one's personal experience (-1 = no, 1 = yes) significantly decreased the beta weight of the reference point condition from $\beta = -.142$, $t(376) = 2.80$, $p = .005$, to $\beta = -.080$, $t(375) = 1.50$, $p = .134$, Sobel test = 3.03, $p = .002$, in support of mediation.

Alternative Explanations. Although we find evidence for our proposed process, we considered the possibility that other differences between goals and expectations could similarly explain the effect. We tested two such possibilities. First, if goals are more vague (e.g., "I want to do well"), people may more easily convince themselves that their goal (vs. expectation) was met. To examine this possibility, the open-ended responses were coded for whether having met the goal/expectation could be objectively judged. Statements were coded as a 1 if it was clear and unambiguous that the goal or expectation was or wasn't met and -1 otherwise. In contrast to the possibility that goals are vaguer, goals were rated as objective in 42% of cases whereas expectations were rated as objective in only 13% of cases, $\chi^2(1) = 39.83$, $p < .001$. Further, objectivity did not predict the likelihood of recommending the course, $\beta = .112$, $t(377) = 1.21$, $p = .228$.

The second alternative explanation is related to the control that people have on meeting goals versus expectations. For example, it is possible that goals are more within the person's

control than their expectations. If this is true, then participants may have been able to exert effort throughout the semester to ensure their goals—but not their expectations—were met. To examine this second possibility, the statements were coded as 1 if the student's (in)actions and attitudes could affect whether the goal/expectation was met (e.g., "Coming to lecture") and -1 if the ability to meet the goal/expectation was outside of a person's control (e.g., "It would be a tough and fast pace course"). Many more of the goals (52%) were rated as being within the student's control compared to the expectations (14%), $\chi^2(1) = 60.39$, $p < .001$. However, whether the statement was within the student's control did not predict the likelihood of recommending the course, $\beta = .043$, $t < 1$, $p = .655$.

Discussion. In sum, study 2 extends the results of study 1 by examining the effect of expectations on the evaluation of college courses, a different type of experiential purchase. Participants were less likely to recommend college courses they had taken when considering their expectations rather than their goals for the course. Finally, study 2 finds additional support for our proposed process explanation. Participants' expectations were less likely to focus on the students' personal experiences in the course relative to goals. This reduced focus on one's personal experience explained the difference in recommendation likelihood. Other differences between goals and expectations, such as controllability and objectivity in judgment, could not similarly explain these results.

We replicate this finding in a third domain: movie watching. In a field study at a movie theater, we demonstrate lower willingness to recommend a movie when expectations versus goals have been evoked (see supplemental study 2 in the web appendix for details).

STUDY 3: DRAWING THE EXPERIENCE

Studies 1 and 2 found that consideration of expectations decreased recommendation likelihood of experiential purchases. Moreover, these studies provided evidence that this effect results from shifting consumers' focus away from one's personal experience toward the experience's attributes. However, it is possible that the difference in focus on one's personal experience through coding of written responses that we have shown thus far is simply a function of the language used to express expectations versus goals rather than a difference in how the experiential purchase is being conceptualized. To rule out this possibility, the next study builds on the earlier studies in two ways.

First, study 3 employs three conditions: expectations, goals, and experience. Although Study 1 included a control condition that was not led to think of a specific reference point, it is possible that goals are the natural reference point used by consumers. If so, the effects of earlier studies could be driven by goals increasing recommendation likelihood. By including a condition that asks participants to describe their experience explicitly, we can provide greater evidence that this effect is due to a decrease in recommendations when considering one's expectations rather than an increase in recommendations when considering one's goals. Moreover, in the study, we ask participants to draw a picture representing either their expectations for, their goals for, or their experience at the restaurant visit before they indicate their likelihood of recommending the restaurant to others. To the extent that expectations change how consumers conceptualize their experiential purchases, we predicted that considering expectations would reduce the likelihood that participants drew themselves (a person) in their drawing of the restaurant visit. Such evidence would provide greater support for our proposed process.

Method

      This was a three-condition study with reference point (control, goals, or expectations) manipulated between subjects. Three hundred and eighty-eight undergraduate students, as determined by the number of available lab participants during the semester, completed this paper and pencil survey in exchange for partial course credit (48% female, M_{age} = 20.73, SD = 2.29).

      We asked participants to think about the last sit-down (non-fast food) restaurant where they had eaten. They first provided the name of the restaurant and the date of the visit. Next, depending on condition, we asked participants to describe their expectations (expectation condition) or goals (goals condition) for visiting the restaurant or to describe their experience at the restaurant (experience condition). Subsequently, we asked people to draw either their goals for visiting, expectations for visiting, or experience at the restaurant.

      As in previous studies, we also asked about participants' likelihood of recommending the restaurant (1 = very unlikely, 7 = very likely). We prefaced this measure with "Considering your goals/expectations/experience…" in the corresponding conditions. Participants then rated their satisfaction with the restaurant visit that was also prefaced with "Considering your goals/expectations/experience…" in the corresponding conditions (1 = very unsatisfied, 7 = very satisfied).

      We then included two measures to assess conscious awareness of the extent to which people thought about their personal experience and emotional response to the experience versus the experience's attributes while determining their willingness to recommend the restaurant. The first measure asked, "When evaluating your willingness to recommend the restaurant, to what extent did you think about restaurant-focused factors (e.g., quality of the chef, décor style, server

knowledge and skills)?" (1 = not at all, 7 = very much). The second measure asked, "When

evaluating your willingness to recommend the restaurant, to what extent did you think back on

your personal reactions to the experience (e.g., thinking about how you felt, imagining

consuming the food)?" (1 = not at all, 7 = very much). We then measured the cost of the meal,

the type of cuisine, whether they had visited this restaurant before, their age and their gender.

Results and Discussion

Recommendations. The overall effect of condition on likelihood of recommending the

restaurant was marginally significant, $F(2, 385) = 2.92$, $p = .055$. More importantly, the planned

contrasts revealed the expected pattern of results. Participants in the expectations condition ($M =$

5.80, SD = 1.25) were significantly less likely to recommend the restaurant than participants in

the other conditions, $F(1, 385) = 5.84$, $p = .016$. Finally, there was no difference in willingness to

recommend the restaurant between participants in the experience condition ($M = 6.09$, SD =

1.07) and those in the goal condition ($M = 6.09$, SD = 0.99), $F < 1$.

Satisfaction. Although the means for satisfaction were in a similar pattern ($M_{expectation} =$

5.90, SD = 1.21; $M_{goal} = 6.06$, SD = 0.99; $M_{experience} = 5.99$; SD = 1.18) to the recommendation

measure, the omnibus test and follow-up contrasts failed to reach significance, all $F < 1$.

Open-ended responses. We first examined participants' open-ended responses for a focus

on their personal experiences. In this study, statements were coded by participants on Mturk.

Responses were coded as focused on one's personal experience if at least half of the Mturk

ratings indicated that the statement referred to one's personal experience using a binary measure[2]

---

[2] Results become stronger if the mean rather than the binary measure is used.

(see web appendix for detailed procedure). We created two orthogonal contrasts to examine the impact of this measure through regressions. The first was a variable contrasting the expectations condition to the goals and experience conditions (expectations = 2, goals = -1, experience = -1). The second was a contrast of the goals condition and the experience condition (expectations = 0, goals = 1, experience = -1). Expectations condition negatively predicted focus on one's personal experience $\beta = -.469$, $\chi^2 = 36.75$, $p < .001$, and focus on one's personal experience positively predicted the likelihood of recommending the restaurant, $\beta = .317$, $t(386) = 2.78$, $p = .006$. When we regressed the likelihood of recommending the restaurant on orthogonal contrasts for condition (expectations: expectations = 2, goals = -1, experience = -1; goals vs. experience: expectations = 0, goals = 1, experience = -1), including the binary measure of focus on one's personal experience significantly decreased the beta weight of the expectations condition contrast from $\beta = -.122$, $t(385) = -2.42$, $p = .016$, to $\beta = -.086$, $t(384) = -1.62$, $p = .106$, Sobel test = -2.07, $p = .038$, in support of mediation.

Drawings. Next, we examined the drawings participants made of their expectations for, goals for, or experience during the restaurant visit to examine whether the same shift in focus would present itself in participants' visual depictions of their restaurant visits. To do so, a research assistant blind to the hypothesis coded participants' drawings for whether a person was included in the drawing (-1 = no, 1 = yes). See figure 3 for samples of drawings. Thirteen participants did not draw anything and were thus excluded from the following analysis. In line with the premise that the inclusion of a person in the drawing relates to people conceptualizing the experience more in terms of their own personal experience, whether a person was included in the drawing was significantly and positively correlated with whether the participants' open-ended response was coded as being focused on their personal experience, r = .221, $p < .001$

(further providing construct validity). Whereas participants drew people only 43.2% of the time in the expectations condition, they drew people 65.5% and 74.8% of the time in the goals condition and experience condition respectively. Indeed, binary logistic regression confirmed that participants in the expectations condition were less likely than participants in the goals and experience conditions to include a person in the drawing, B = -0.380, SE = 0.076, Wald χ^2 = -24.96, $p < .001$. The difference between the goals condition and the experience condition was not significant, B = -0.223, SE = 0.139, Wald χ^2 = -2.55, $p = .110$.
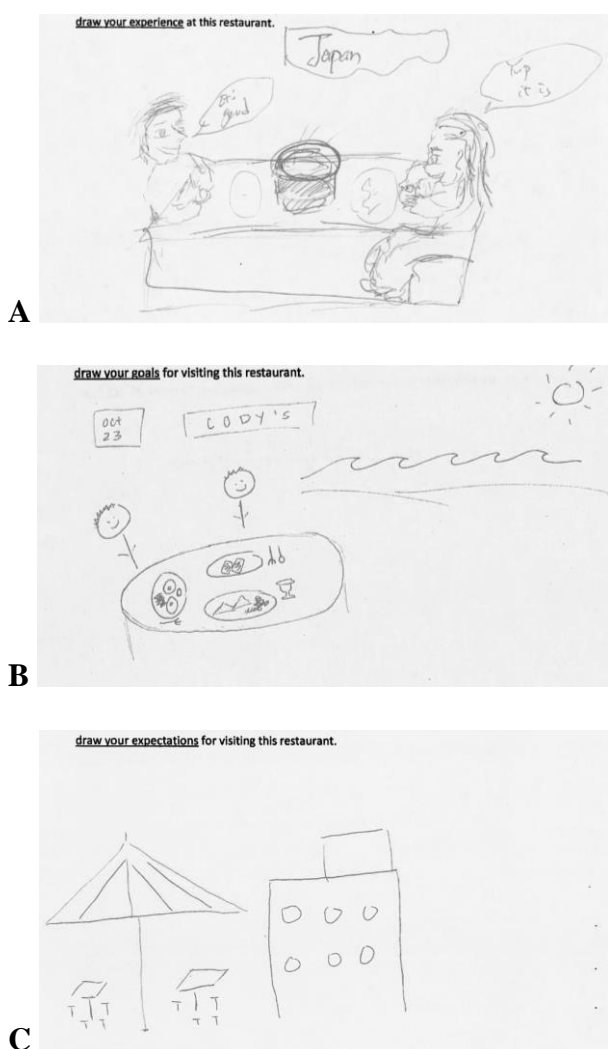

FIGURE 3. Sample Drawings (Study 3).
Note: Panel A/B/C depict sample drawings from the experience/goal/expectations conditions.

We next examined the relationship between participants' inclusion of a person in their drawings and their likelihood of recommending the restaurant. Regressing participants' inclusion of a person in their drawing of the restaurant visit significantly and positively predicted participants' likelihood of recommending the restaurant, B = 0.24, SE = 0.12, $t(373) = 2.03$, $p = .043$. When we regressed the change in likelihood of recommending the restaurant on reference point conditions (using the orthogonal contrasts), including whether people were included in the drawings (1 = yes, 0 = no) directionally decreased the beta weight of the expectations condition contrast from $\beta = -.138$, $t(372) = -2.70$, $p = .007$, to $\beta = -.120$, $t(371) = -2.25$, $p = .025$, though this difference did not reach significance, Sobel test = -1.30, $p = .194$.

Self-reported focus. There were no differences by condition in self-reports of relying on restaurant-focused factors or of personal reactions to the experience, all $F < 1$. Given the differences in the coding of the written responses as well as the differences in the drawings people created, these results suggest that either consumers are not consciously aware of the factors they consider when deciding to recommend experiential purchases or that the self-report measures were not sensitive enough to capture these differences.

Discussion. This study replicates and extends the findings of the previous studies. Participants asked to consider their expectations were less willing to recommend the restaurant than other participants. Further, this study provides evidence for the shift in focus away from one's personal experience in a novel form. Drawings of restaurant visits were less likely to include people when thought about as a function of expectations compared to goals or one's experience. These findings provide further evidence that expectations change how one conceptualizes an experiential purchase such that consumers focus less on their own personal experience and emotional reactions.

STUDY 4: EXPERIENCES VERSUS PRODUCTS

Studies 1-3 showed that thinking about expectations reduces the likelihood of recommending experiential purchases. Moreover, we provided evidence that this effect occurs because expectations draw attention away from one's own experience and emotional reactions. However, an alternative possibility is that expectations are consistently too high and thus reflecting on expectations will dampen consumers' willingness to recommend anything.

To differentiate these two possibilities, in study 4, in addition to varying the consideration of expectations, we vary whether participants consider an experiential purchase or a material purchase. Compared to experiential purchases that are made with the purpose of living through (to get a personal experience), material goods are tangible purchases kept in one's possession. As such, when considering recommending a material purchase, consumers should more readily think about the purchase's attributes and reflect less on their personal experience(s) to begin with. If the effect of expectations is simply a function of expectations being too high, expectations should decrease consumers' willingness to recommend both experiential and material purchases alike. However, if as we suggest, the decrease in recommendations results from shifting the focus away from one's personal experience and emotional reactions, consumers' reduced willingness to recommend resulting from the consideration of expectations should be specific to experiential purchases.

Method

Participants in the main study were 400 people from Mechanical Turk who completed the study in exchange for monetary compensation. The study followed a 2 (reference point: control vs. expectations) X 2 (purchase: experience vs. material good) between-subjects design.

Participants were first asked to recall a recent purchase. Half of the participants recalled an experiential purchase (an event that they paid for and recently went to) and the other half recalled a material purchase (a discretionary item that they purchased and recently received). All participants provided a brief description of their purchase. Next, we asked participants in the expectations condition to reflect on and describe their expectations for the event or item while participants in the control condition did not complete this task. We then asked participants about their willingness to recommend the purchase with those in the expectations condition explicitly asked to consider their expectations when making this decision. First, participants saw the following question: "Thinking about [your expectations for] this [event/item], how likely would you be to recommend this [event/item] to someone you know? (1 = definitely not recommend, 7 = definitely recommend; wording in parenthesis varied by condition). Additionally, we asked how they would rate the purchase if they were to review it (1 - 10 stars).

Participants then completed a page with demographic questions (age, gender, and language) as well as an instructional manipulation check (Oppenheimer, Meyvis, and Davidenko 2009; see web appendix for details). Sixteen participants failed the instructional manipulation check and were excluded from all analyses, leaving a final sample of 384 participants (M_{age} = 35.66, SD = 11.72, 48.2% female).

Results and Discussion

Willingness to Recommend. There was no main effect of reference point or purchase type, both $F \leq 1.01$, $p \geq .317$. However, there was a significant reference point by purchase type interaction, $F(1, 380) = 12.81$, $p < .001$. Follow-up contrasts were in line with our predictions.

Replicating earlier studies, expectations significantly reduced recommendation likelihood when the purchase was experiential ($M_{expectation}$ = 5.73, SD = 1.53 vs. $M_{control}$ = 6.11; SD = 1.16), $F(1, 380)$ = 3.92, p = .049. However, expectations significantly *increased* recommendation likelihood when the purchase was material ($M_{expectation}$ = 6.07, SD = 1.18 vs. $M_{control}$ = 5.50; SD = 1.29), $F(1, 380)$ = 9.59, p = .002). See figure 4.
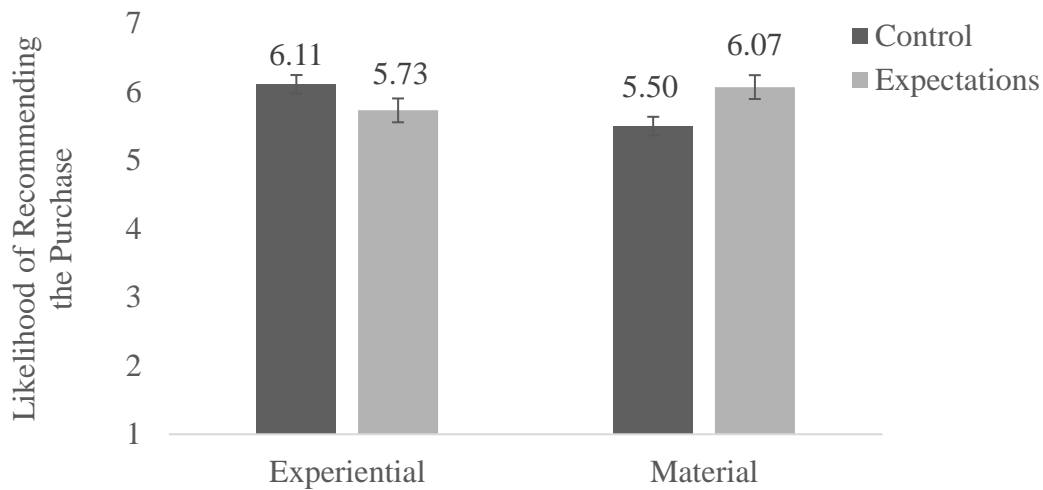

FIGURE 4. Likelihood of Recommending the Purchase by Reference Point Condition and Purchase Type (Study 4).

Note: Error bars represent standard error of the mean.

Star Ratings. Replicating the recommendation measure, there was no main effect of reference point or purchase type, both F < 1, but there was a significant reference point by purchase type interaction, $F(1, 380)$ = 5.52, p = .019. Although directionally consistent, the difference in ratings of the experiential purchase by reference point did not reach significance ($M_{expectation}$ = 7.97, SD = 1.88 vs. $M_{control}$ = 8.33; SD = 1.72), $F(1, 380)$ = 2.23, p = .136. Participants' ratings of their material goods were marginally higher when using their

expectations as the reference point ($M_{expectation}$ = 8.45, SD = 1.41 vs. $M_{control}$ = 8.02; SD = 1.62), $F(1, 380) = 3.36$, $p = .067$).

Discussion. Study 4 demonstrates that the effect of expectations on reductions in willingness to recommend is unique to experiential purchases. In fact, expectations increased participants' likelihood of recommending their material purchases. Although the reversal was unexpected, this data casts doubt on the alternative explanation that expectations are simply an unrealistically high reference point and therefore unlikely to be met. Instead, this study is consistent with the premise that people typically focus on their own personal experience when deciding whether to recommend an experiential purchase and that expectations shift consumers' focus toward the experience's attributes, which reduces recommendation likelihood.

STUDY 5: TRIPADVISOR REVIEWS

The previous studies demonstrate the impact of expectations on recommendation likelihood across a variety of experiential purchases and a combination of field and lab studies. In this next study, we use a large sample of real customer reviews collected from a popular review platform, TripAdvisor, to provide greater external validity, to replicate our effect with another type of experiential purchase (hotel stay), and to demonstrate that the natural occurrence of expectations being used as a reference point produces a similar effect.

Data

Launched in 2000, TripAdvisor is one of the most popular review platforms for travel related content. During 2016, TripAdvisor reached 500 million customer reviews and opinions

about over 6 million accommodations, restaurants, and attractions.[3] We collected the entire

review history of 7,874 hotels in the state of California. The review sample contains 1,931,859

reviews written between 2001 and 2016. For every review, we obtained its unique ID, the date of

publication, the text content, and its star-rating. Moreover, for every review, we recorded the

unique ID of the user who wrote the review and the month and year in which they stayed at the

hotel reviewed.[4] Finally, for all reviewers who reviewed a hotel in our dataset, we collected their

entire review history, for a total of about 23 million reviews from 1.2 million reviewers.

Review Content Coding

Expectations. Using text search, we identified reviews for which one's expectations were

used as the reference point. Specifically, we coded each review as using expectations as the

reference point if the review contained any variation of the word "expect" (e.g., expectation,

expected, unexpected) at least once. We manually checked 100 randomly chosen reviews to

make sure that the coding was correct.

Focus on One's Personal Experience. Our proposed process suggests that consumers

considering their expectations should be more focused on external factors and less focused on

their own personal experience and emotional reactions. To examine this account in the online

reviews, we measured the extent to which reviews are emotional using LIWC (Pennebaker et al.

2015), a software for automated text analysis. Specifically, we measured the percentage of words

contained in the review that were affective in nature, either positive (e.g., happy, love) or

[3] See: https://www.tripadvisor.nl/pages/factsheet.html
[4] Even though this date information is not mandatory, the percentage of consumers disclosing it is very high. In our data, 98.7% of reviews include the year-month of stay at the hotel.

negative (e.g., annoyed, hate). If expectations shift consumers' focus from their own personal experience and emotional reactions to more external factors, as we predict they do, consumers using their expectations as the reference point should include fewer affective words in their reviews. Further, such reduction in affective words should help explain the effect of expectations on evaluations. Summary statistics of the TripAdvisor dataset are available in table 1 of the web appendix.

Model Specification and Results

The review's star-rating served as the dependent measure.[5] To test whether thinking about expectations affected consumers' likelihood of recommending a hotel, we used the following specification:

$$Stars_{ijt} = \beta\ Expectation_{ijt} + \alpha\ Affect_{ijt} + X_{ijt}\gamma + \alpha_j + \tau_t + \epsilon_{ijt} \qquad (1)$$

where the dependent variable is the star-rating of review i of hotel j written in year- month t. $Expectation_{ijt}$ and $Affect_{ijt}$, whose coefficients are of interest, are an indicator of whether the review i of hotel j written in year-month t is identified as describing expectations, and the LIWC measure of affective words (i.e., the percentage of affective words) for review i of hotel j written in year-month t. α_j and τ_t are hotel and year-month fixed effects, respectively. Hotel fixed effects help control for time-invariant unobservable hotel characteristics that may affect ratings, while time fixed effect controls for shocks to ratings common to all the hotels. X_{ijt} is a vector of controls in which we include review environment characteristics such as the logarithm of the

---

[5] While this measure can be conceptualized as a consumer's evaluation, since it is posted online, it is likely meant for the purposes of influencing others and hence can be seen as a recommendation. Indeed, in the web appendix, we provide evidence that star-ratings are a good proxy for willingness to recommend the hotel.

number of reviews prior to review i and the hotel average ratings prior to review i. Finally, following standard practices, we cluster standard errors at the individual hotel level (Bertrand, Duflo, and Mullainathan 2004; Donald and Lang, 2007).

We report the results of this specification in table 1. In the first column, we present the estimates without the variable $Affect_{ijt}$ and without any controls. The coefficient of interest, $Expectation_{ijt}$, is negative and statistically significant, suggesting that when consumers consider their expectations, their star-rating is, on average, .2-stars lower.

Table 1: The Effect of Expectations on Review Star-Ratings: Hotel Fixed Effects

	(1)	(2)	(3)
Expectations	-0.227^{***}	-0.172^{***}	-0.172^{***}
	(0.005)	(0.005)	(0.005)
Affect		0.072^{***}	0.072^{***}
		(0.001)	(0.001)
Avg. Stars$_{t-1}$			0.293^{***}
			(0.009)
log Reviews$_{t-1}$			-0.018^{***}
			(0.004)
N	1931556	1931556	1923769
R^2 within	0.0038	0.076	0.079

Note: The dependent variable is the star-rating of review i of hotel j written at time t. Cluster-robust standard errors at the individual hotel level are shown in parentheses.
Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Notably, because most review platforms round their ratings to the nearest half-star, a .2 difference can be meaningful. For example, on TripAdvisor, a hotel with an average star-rating of 3.4 will have a displayed star-rating of 3.5 stars. However, a hotel with 3.2-stars average (only 0.2 stars lower) will have a displayed star-rating of 3-stars. Moreover, such half-star differences

can generate substantial loss for the firms. For example, Luca (2016) shows that, for restaurants reviewed on Yelp, a half-star increase in displayed star-ratings can generate up to 4.5% increase in revenue.

Next, we examine the relationship between the use of expectations and the amount of affective words used in the review. In line with predictions, reviews using expectations as a reference point used fewer affective words ($\beta = -.76,\ SE = .015, p < .001$).

In column 2 of table 1, we incorporate the affective words variable to test our proposed process. The results confirm the process observed in our lab and field experiments: affective words positively predict ratings, and the effect of expectations decreases in magnitude by about 23%.[6] This result suggests that the effect of expectations can be in part explained by a reduction in using affective words to evaluate the hotel and is in line with our hypothesis that expectations shift consumers' focus away from one's personal experience and emotional reactions. Finally, in column 3, we test the robustness of the results by including review environment controls. Again, we observe a negative and statistically significant coefficient very similar in magnitude to our previous estimate. In the web appendix, we show that these results are robust to the inclusion of hotel-year-month of stay fixed effects, which allow us to compare reviewers that stayed at the *same* hotel at the *same* time.

Exploiting the Reviewer ID

In the previous analysis, we showed that reviews describing expectations have, on average, lower ratings than the rest of the reviews. Arguably, the key concern with the preceding

---

[6] A Hausman-type cross-model test rejects the hypothesis that the coefficients of Expectations in columns 1 and 2 of table 1 are the same.

identification is that our results could be driven by differences in the people writing the reviews for the same hotel. That is, it is possible that consumers who use expectations in their reviews might be *different* from those that do not. For instance, people who are generally more negative in their evaluations may be more likely to use expectations in their reviews. To examine this possibility, we take advantage of the richness of the TripAdvisor dataset and estimate a specification that includes hotel-reviewer fixed effects. By doing so, the impact of expectations is identified by changes in star-ratings of reviews that are about expectations versus those that are not, written by the *same* reviewer for the *same* hotel. The specification we estimate takes the following form:

$$Stars_{ijkt} = \beta\, Expectation_{ijkt} + +\alpha\, Affect_{ijkt} + \gamma X_{ijkt} + \eta_k \times \alpha_j + \tau_t + \epsilon_{ijkt,} \qquad (2)$$

where the dependent variable is the star-rating of review i of hotel j written by reviewer k in year-month t. $Expectation_{ijkt}$ is an indicator of whether review i is identified as describing expectations, and $Affect_{ijkt}$ is the percentage of affective word of review i. $\eta_k \times \alpha_j$ are the hotel – reviewer fixed effects. In X_{ijkt} we include controls such as the average ratings of the reviews written by reviewer j before review i, and the total number of reviews written by reviewer j before review i.[7] Finally, following standard practices, we cluster standard errors at the individual hotel-user level. We report the results of such specification in Table 2. As before, in column 1 we present the results without the variable $Affect_{ijkt}$ and without any controls. The coefficient of interest, $Expectations_{ijkt}$, is negative and significant, suggesting that a reviewer

---

[7] We also tested a specification in which we included controls for the review environment (i.e., the logarithm of the number of reviews prior to review *i* and the hotel average ratings prior to review *i*) and obtained similar results.

who evaluates the same hotel multiple times will rate the hotel as significantly worse if they were considering their expectations at the time of making the review. In column 2, we add the affective words variable to the model. As expected, the coefficient of $Expectations_{ijkt}$ decreases in magnitude, while $Affect_{ijkt}$ is positive and statistically significant. Finally, in column 3, we include the vector of controls, and obtain similar results.

Table 2: The Effect of Expectations on Review Star-Ratings: Reviewer-Hotel Fixed Effects

	(1)	(2)	(3)
Expectations	-0.080^{***}	-0.058^{***}	-0.056^{***}
	(0.016)	(0.015)	(0.015)
Affect		0.033^{***}	0.031^{***}
		(0.001)	(0.001)
User has Previous Reviews			1.101^{***}
			(0.050)
User has Previous Reviews $\times$ Avg. User Stars$_{t-1}$			-0.235^{***}
			(0.010)
log User Reviews$_{t-1}$			-0.031^{***}
			(0.006)
N	70975	70975	68214
R^2	0.00093	0.029	0.055

Note: The dependent variable is the star-rating of review i of hotel j written by reviewer k at time t. Cluster-robust standard errors and at the individual hotel-user level are shown in parentheses.
Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Conclusion

Overall, using millions of reviews on TripAdvisor, we replicate the negative impact of using expectations as a reference point on consumers' recommendations of their experiential purchases. These results were robust to different specifications that rely on increasingly stringent identifications, and could not be explained by differences in the types of hotels reviewed or by

differences in the types of people writing reviews while thinking about expectations. Further, consumers considering their expectations tended to write less about their emotional reactions to the experience in line with the premise that expectations used as reference points diverts attention away from one's personal experience and emotional reactions towards the more external attributes of the experience offering. Indeed, we demonstrated that the dampened evaluations by consumers considering their expectation was in part explained by reduced affective content included in their reviews.

STUDY 6: EXPEDIA REVIEWS

Next, we sought to replicate our effect in yet another large dataset. Thus, we collected hotel data from Expedia, another popular hotel reviewing site. By mid-2017, the company provided reviews for over 380,000 accommodation properties and declared 275 million room nights booked.[8] This site differs from TripAdvisor in a couple of important ways. First, consumers can only review on Expedia if they purchased their stay through Expedia. Thus, all reviews are verified purchases. Second, Expedia provides a different operationalization of consumers' evaluations. More closely aligned with our primary dependent measure throughout our earlier studies, Expedia asks consumers whether they would recommend the hotel to others. It then aggregates this information into a percentage of consumers willing to recommend each individual hotel. Thus, in this dataset, we examine whether the proportion of consumers who consider their expectations affects the proportion of consumers willing to recommend the hotel.

---

[8] See: http://www.expediainc.com/about/

Data

We collected the entire review history of all U.S. hotels listed on Expedia with at least one review. This dataset contains over 12 million reviews from 39,672 hotels written between 2004 and 2015. For every review, we recorded the date of publication, the star-rating associated with the review, and the text content. Moreover, for every hotel, we collected information regarding the percentage of consumers willing to recommend the hotel as indicated by a single aggregate value for each hotel.

Review Content Coding

As in study 5, we coded each review as using expectations as the reference point if the review contained any version of the word "expect." We calculated the percentage of reviews for each hotel that considered expectations. Additionally, we used LIWC to measure the percentage of words contained in each review that were affective in nature. We then aggregated this measure at the hotel level by averaging this number across each hotel's reviews.[9] Summary statistics are available in table 1 of the web appendix.

Model Specification and Results

We estimate the impact of expectations on recommendations using the following model:

$$Pct\ Recommendations_i = \beta\ Pct.Expectations_i + \alpha\ Affect_i + X_i\gamma + \epsilon_i, \quad (3)$$

---

[9] The aggregation at the hotel level of Affect and Expectation is necessary because the recommendation measure obtained from Expedia is at the hotel level. Thus, differently from the TripAdvisor analysis that exploited the panel nature of the dataset, the Expedia analysis is cross-sectional.

where the dependent variable is the percentage of reviewers recommending hotel i as indicated

by Expedia., $Pct. Expectations_i$ is the percentage of hotel reviews identified as describing

expectations, and $Affect_i$ is the average use of affective words in the reviews of hotel i. Finally,

X_i is a vector of controls in which we include a hotel city dummy to account for differences in

the percentage of recommendations due to the hotel location that might be correlated with the

review content. We report the estimates of this specification in table 3.

Table 3: The Effect of Expectation on Hotel Recommendations

	(1)	(2)	(3)
Pct. Reviews About Expectations	−0.228***	−0.134***	−0.117**
	(0.047)	(0.047)	(0.049)
Affect		2.471***	2.505***
		(0.145)	(0.169)
City Dummy	No	No	Yes
N	39672	39672	38053
R^2	0.0022	0.069	0.19

Note: The dependent variable is the percentage of users recommending hotel i. Robust standard errors are shown in parentheses.
Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

In column 1, we present the results without the variable $Affect_i$ and without any

controls. As expected, the coefficient of interest is negative and statistically significant,

suggesting that a 10 percentage point increase in the percentage of reviewers considering their

expectations while providing their review generates a decrease in recommendations of

approximately 2 percentage points. In column 2, we include the measure of affective words and

observe that affective words positively impact ratings. Moreover, in support of the proposed

process, including affect in the model significantly decreases the magnitude of the effect of

expectations on recommendations.[10] Finally, in column 3, we include the city dummy and obtain

very similar estimates.

Conclusion

 This archival data provides further convergent evidence of the proposed effect and

underlying mechanism. Of course, the evidence from this study and the previous study are

correlational in nature, and we considered the possibility that the effects found here are due to

reverse-causality—that people are more likely to spontaneously consider and write about

expectations when they are not met. If this were the case, reviews of all types should be

worsened when consumers reflect on their expectations. However, against this possibility, in

analysis of reviews from Amazon.com, we found no evidence that expectations decrease

recommendations for material goods (see supplemental study 3 in the web appendix).

GENERAL DISCUSSION

 Recommendations are incredibly important to companies' growth and future sales

(Reichheld 2003). This is particularly true for experiential purchases because the quality of

experiential purchases is often otherwise difficult to determine prior to consumption (e.g., Litvin,

Goldsmith, and Pan 2008). This research is among the first to consider how horizontally

differentiated reference points may influence consumers' evaluations. We demonstrate that the

use of expectations as a reference point systematically decreases consumers' likelihood of

---

[10] A Hausman-type cross-model test rejects the hypothesis that the coefficients of Expectations in columns 1 and 2 of table 3 are the same.

recommending an experiential purchase. This effect emerges in field studies, lab studies, as well as in analyses of millions of online reviews.

Our results cannot be explained by differences in one's actual experience, as the result emerges when expectations are made salient *after* an experience is over (studies 1-4), and even when they are made salient after consumers make initial evaluations (study 1). Our results demonstrate that the use of expectations as one's reference point changes how consumers conceptualize their experience. Expectations shift consumers' focus away from their own personal experience and emotional responses toward the experience's attributes. Evidence for this proposed process is demonstrated by a variety of means. Expectations resulted in a reduced likelihood of elaborating on one's own personal experience as coded by research assistants (studies 1 and 2), crowdsourcing (study 3), and text analysis software (studies 5 and 6). Additionally, this difference extended to how people drew an experience: people considering their expectations were less likely to draw themselves when asked to provide a visual representation (study 3). Moreover, the decrease in recommendation likelihood did not extend to material goods, purchases that are less likely to be conceptualized as a personal experience to begin with (study 4).

We note that the impact of expectations on satisfaction ratings was consistently weaker than its impact on recommendations and often non-significant. Indeed, our explanation for why this effect occurs provides reason to believe satisfaction may not be affected by expectations to the same extent. First, we build our theory on the drivers of word of mouth, which examines what motivates people to engage in word of mouth. Thus, the effect is likely to be stronger for dependent measures related to the decision to engage in word of mouth compared to measures related to evaluations more broadly. Additionally, evaluations such as satisfaction may be more

readily conceptualized as one's emotional response to an experiential purchase, though future research is necessary to better understand these potential differences.

Indeed, this research offers a number of new directions for future research. Research finds that expectations are good predictors of evaluations (e.g., Johnson, Anderson, and Fornell 1995). However, the current findings suggest that consumers are not naturally, explicitly considering their expectations when they decide whether to recommend a range of experiential purchases. Future research should try to better understand this seeming paradox. That is, why are expectations good predictors of evaluations if they are not being used as the reference point? For instance, to the extent that expectations formed prior to an experience reflect one's best estimate of what their experience will be like, perhaps expectations are a means of capturing individual differences in preferences toward types of experiences (e.g., people who like Italian food have higher expectations for Italian restaurants). Alternatively, perhaps expectations lead to different considerations depending on whether they are considered implicitly versus explicitly.

Another open question is whether recommendations resulting from the consideration of expectations are more or less accurate than those resulting from other types of reference points. That is, although expectations decrease recommendations overall, to the extent that expectations focus consumers on more time invariant attributes of the experience offering, do they more accurately capture what someone else is likely to experience another time? Relatedly, how do consumers listening to or reading other people's recommendations respond to recommendations focused on attributes of the experience offering versus assessments of one's personal experience? Is one type more persuasive than the other? Finally, although this research is focused on recommendations of experiential purchases, it is worth noting the unexpected reversal in study 4 for material goods. This study found that expectations actually increased consumers' likelihood

of recommending their material goods. As this research is among the first to examine how horizontally differentiated reference points can systemically impact evaluations, future research would benefit from a better understanding of how and why different reference points influence different types of purchases and methods of evaluations.

Our results have important managerial implications and should serve as a guide to companies eliciting reviews from consumers or encouraging word-of-mouth campaigns. First, companies would be wise not to make expectations salient when asking consumers to recommend an experiential purchase to others. As such, companies should not ask consumers to disclose their expectations or to rate their experience as a function of their expectations. Moreover, the underlying process we identify provides implications beyond the effect of expectations specifically. Our results suggest that any factor that shifts one's focus toward or away from one's personal experience and emotional reactions will impact recommendation likelihood. For example, end of experience surveys asking people to evaluate aspects of the experience's attributes may dampen recommendation likelihood. By contrast, consumers may be more likely to recommend experiential purchases to others if they are focused on the experience – for instance, by prompting people to share their experience with others.

REFERENCES

Anderson, Rolph E. (1973), "Consumer Dissatisfaction: The Effect of Disconfirmed Expectancy on Perceived Product Performance," *Journal of Marketing Research*, 10 (1), 38-44.

Baumgartner, Hans, Mita Sujan, and James R. Bettman (1992), "Autobiographical memories, affect, and consumer information processing." *Journal of Consumer Psychology* 1, 1, 53-82.

Berger, Jonah, (2014), "Word of mouth and interpersonal communication: A review and directions for future research." *Journal of Consumer Psychology* 24, 4, 586-607.

Berger, Jonah, & Katherine L. Milkman (2012), "What makes online content viral?" *Journal of Marketing Research*, *49*(2), 192-205.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan (2004), "How Much Should We Trust Differences-in-Differences Estimates?" *The Quarterly Journal of Economics*, 119 (1), 249-275.

Carter, Travis J., and Thomas Gilovich (2010), "The relative relativity of material and experiential purchases." *Journal of Personality and Social Psychology,* 98, 1, 146.

Chung, Cindy MY, and Peter R. Darke (2006), "The consumer as advocate: Self-relevance, culture, and word-of-mouth." *Marketing Letters* 17, 4, 269-279.

Deighton, John (1984), "The Interaction of Advertising and Evidence," *The Journal of Consumer Research*, 11 (3), 763-70.

Donald, Stephen G. and Kevin Lang (2007), "Inference with Difference-in-Differences and Other Panel Data," *The Review of Economics and Statistics,* 89 (2), 221-233.

Hoch, Stephen J. and Young-Won Ha (1986), "Consumer Learning: Advertising and the Ambiguity of Product Experience," *Journal of Consumer Research*, 13 (2), 221– 33.

Johnson, Michael D., Eugene W. Anderson, and Claes Fornell (1995), "Rational and Adaptive Performance Expectations in a Customer Satisfaction Framework," *Journal of Consumer Research*, 21 (4), 695-707.

Kahneman, Daniel (1992), "Reference points, anchors, norms, and mixed feelings," *Organizational behavior and human decision processes* 51.2 (1992): 296-312.

Kahneman, Daniel, Barbara L. Fredrickson, Charles A. Schreiber, and Donald A. Redelmeier (1993), "When More Pain is Preferred to Less: Adding a Better End," *Psychological Science*, 4 (6), 401-405.

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review, 93*(2), 136-153.Kahneman, Daniel and Amos Tversky (1979), "Prospect Theory: An Analysis of Decision under Risk," Econometrica. 47 (2): 263.

Kumar, V., J. Andrew Petersen, and Robert P. Leone (2007), "How Valuable is Word of Mouth?" *Harvard Business Review*, 85 (10), 139.

Langston, Christopher A. (1994), "Capitalizing on and coping with daily-life events: Expressive responses to positive events." *Journal of Personality and Social Psychology* 67, 6, 1112.

Litvin, Stephen W., Ronald E. Goldsmith, and Bing Pan (2008), "Electronic Word-of-Mouth in Hospitality and Tourism Management," *Tourism Management*, 29 (3), 458–468.

Liu, Yong (2006), "Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue," *Journal of Marketing*, 70 (3), 74-89.

Luca, Michael (2016), *Reviews, Reputation, and Revenue: The Case of Yelp.com*, Harvard Business School Working Paper.

Neilsen (2014), "#Twothumbsup: Moviegoing at a Theater Near You," retrieved from

    http://www.nielsen.com/us/en/insights/news/2014/twothumbsup-moviegoing-at-a-

    theater-near-you.html.

Ofir, Chezy and Itamar Simonson (2001), "In Search of Negative Customer Feedback: The

    Effect of Expecting to Evaluate on Satisfaction Evaluations," *Journal of Marketing*

    Research, 38 (May), 170-82.

Ofir, Chezy and Itamar Simonson (2007), "The Effect of Stating Expectations on Customer

    Satisfaction and Shopping Experience," *Journal of Marketing Research*, 44 (February),

    164-74.

Oliver, Richard L. (1980), "A Cognitive Model of the Antecedents and Consequences of

    Satisfaction Decisions," *Journal of Marketing Research*, 17 (3), 460-69.

Oliver, Richard L. (1996), *Satisfaction: A Behavioral Perspective on the Consumer*, New York,

    NY: McGraw-Hill.

Oppenheimer, Daniel, Tom Meyvis, and Nicolas Davidenko (2009), "Instructional Manipulation

    Checks: Detecting Satisficing to Increase Statistical Power," *Journal of Experimental*

    Social Psychology, 45 (July), 867–72.

Ordonez, Lisa D., Terry Connolly, and Richard Coughlan (2000), "Multiple Reference Points in

    Satisfaction and Fairness Assessment," *Journal of Behavioral Decision Making*, 13 (3),

    329.

Pennebaker, James W., Ryan L. Boyd, Kayla Jordan, and Kate Blackburn (2015), *The*

    Development and Psychometric Properties of LIWC2015, Austin, TX: University of

    Texas at Austin.

Peters, Kim, Yoshihisa Kashima, and Anna Clark (2009), "Talking about others: Emotionality and the dissemination of social information," *European Journal of Social Psychology* 39, 2, 207-222.

Redelmeier, Donald and Daniel Kahneman (1996), "Patients' Memories of Painful Medical Treatments: Real-Time and Retrospective Evaluations of Two Minimally Invasive Procedures," *Pain,* 66 (1)*,* 3-8.

Reichheld, Frederick F. (2003), "The One Number You Need to Grow," *Harvard Business Review*, 81 (December), 46–54.

Rui, Huaxia, Yizao Liu, and Andrew Whinston (2013), "Whose and What Chatter Matters? The Effect of Tweets on Movie Sales," *Decision Support Systems*, 55, 863-70.

Sujan, Mita, James R. Bettman, and Hans Baumgartner. "Influencing consumer judgments using autobiographical memories: A self-referencing perspective." *Journal of Marketing Research* (1993): 422-436.

Tulving, Endel. "Episodic memory: From mind to brain." *Annual review of psychology* 53.1 (2002): 1-25.

Tuttle, Brad (2013), "Why Some Restaurants Hate Restaurant Week." Retrieved from http://business.time.com/2013/02/28/why-some-restaurants-hate-restaurant-week/.

Van Boven, Leaf, and Thomas Gilovich (2003), "To do or to have? That is the question." *Journal of Personality and Social Psychology,* 85, 6, 1193.

Walker, W. Richard, John J. Skowronski, Jeffrey A. Gibbons, Rodney J. Vogl, and Timothy D. Ritchie (2009), "Why people rehearse their memories: Frequency of use and relations to the intensity of emotions associated with autobiographical memories," *Memory,* 17, 7, 760-773.

Wojnicki, Andrea C., and David Godes (2017), "Signaling Success: Word of Mouth as Self-Enhancement." *Customer Needs and Solutions,* 4, 4, 68-82.

Zwick, Rami, Rik Pieters, and Hans Baumgartner. "On the practical significance of hindsight bias: The case of the expectancy-disconfirmation model of consumer satisfaction." *Organizational behavior and human decision processes* 64.1 (1995): 103-117.

WEB APPENDIX

This web appendix contains two supplemental field studies, one supplemental study of Amazon.com data, details of the baseline measures collected for Study 1, details of the coding of written responses used in Study 3, and supplemental analyses for the TripAdvisor data used in Study 5 of the main paper.

SUPPLEMENTAL STUDY 1: RESTAURANT RECOMMENDATIONS

Customers at a pub were asked to fill out a short customer survey as they paid for their meal. The survey varied whether customers were asked to consider their expectations or their experience at the restaurant prior to indicating their likelihood of recommending the restaurant to others. We predicted that those asked to consider their expectations would be less likely to recommend the restaurant to others compared to those considering their experience.

Method

We partnered with a local pub to conduct this study. The study employed a single-factor (reference point: expectation, control) design with participants randomly assigned across the two conditions. We asked the restaurant to collect approximately 250 responses. Two hundred and forty eight restaurant patrons completed this study. No demographic information was collected. No participants were excluded in this or any other survey reported herein except as otherwise noted.

The survey was a small paper survey included with the check and given to customers by their server at the end of the meal. The restaurant's floor manager received $1 for every completed survey response. We manipulated the reference point by asking some participants to consider their expectations before providing their evaluations. Half the participants (those in the expectations condition), read: "What were your expectations for today's restaurant visit?" The remaining participants (those in the control condition) read: "What was your experience during today's restaurant visit?" All participants were provided with a small area to write open-ended text.

The primary dependent measure was restaurant patrons' likelihood of recommending the restaurant to a friend. Specifically, those in the expectations [control] condition answered, "Considering your expectations [experience], how likely are you to recommend the restaurant to a friend?" (1 = not at all likely, 7 = very likely) as the dependent measure. Participants also indicated whether they had visited the restaurant before. This measure did not affect any of the results and is not discussed further.

Results and Discussion

Recommendation likelihood. Despite near ceiling effects on the dependent measure, in line with predictions, participants in the expectations condition were less likely to recommend the restaurant to others than participants in the control condition ($M_{expectation} = 6.51$, SD = 0.60 vs. $M_{control} = 6.69$, SD = 0.79; $F(1, 246) = 4.03$, $p = .046$).

Open-ended responses. Three coders blind to the hypothesis coded participants open-ended responses using a binary measure indicating whether the statement included the

participants' personal experience. That is, statements that include the self's actions, responses, or how they felt (e.g., to try a new restaurant, to eat good food, enjoy a night out) were coded as a 1 whereas statements that did not include the self (e.g., good value, good food, great server, nice ambiance) were coded as a 0. The coders responses cohered (Cronbach's $\alpha = .88$) and were averaged to form a measure of whether personal experience was included. There were no significant differences on this measure, $F < 1$. Although this measure did not provide support for our hypothesis, the modal length of this written response was one word. Thus, the short responses may have made it difficult to find differences.

This study provides a conceptual replication of the field experiment in study 1 of the main manuscript. When expectations were evoked prior to evaluations, likelihood of recommendation was dampened. This occurred despite finding near ceiling effects on the dependent measure. We speculate that these particularly high responses may have been a function of customers knowing that the survey would be seen by their server. However, as restaurants often include response feedback in this format, the survey format provides ecological validity.

SUPPLEMENTAL STUDY 2: RECOMMENDING A MOVIE

We use a field study to examine the impact of considering expectations on willingness to recommend a movie. Actual patrons at a movie theatre were asked to evaluate a movie after watching it, considering either their goals or expectations for watching the movie.

Method

We offered 100 moviegoers a $4 gift card to complete a brief survey after watching a movie (the gift card was redeemable at that theater). Two participants did not complete the survey and were excluded; all responses are for the remaining 98 participants (average age = 51 years, 61% female). We manipulated the reference point participants considered by having them read a short statement before providing their responses. Half the participants (those in the goal condition), read: "people have different goals for watching movies. Some people watch movies to alleviate boredom, some people watch movies for pleasure, and some others even watch movies for enrichment." The remaining participants (those in the expectation frame) read: "people have different expectations when watching a movie. Some movies fall short of expectations, other movies meet expectations, and some others even exceed expectations." In summary, the study employed a single-factor (frame: goal, expectation) design with participants randomly assigned across the two conditions.

The primary dependent measure was the movie patron's likelihood of recommending the movie to a friend (1 = not at all likely, 7 = very likely). We also asked participants to indicate whether they will recommend this movie (yes/no/don't know).

In addition to measuring their recommendation likelihood, we asked a couple of questions assessing their satisfaction. Patrons reported their satisfaction with their experience of the movie, and their satisfaction with their decision to see the movie (both 7-point scales: 1 = not at all satisfied, 7 = very satisfied), and whether watching the movie was an overall positive experience for them (yes/no/don't know).

Results and Discussion

Recommendation likelihood. We first examined the scale measure of recommendation likelihood. In line with hypotheses, participants in the expectation condition reported a significantly lower likelihood of recommending the movie than those in the goal condition (M_{goal} = 6.31, SD = 1.21 vs. $M_{expectation}$ = 5.52, SD = 2.07); F (1, 96) = 5.30, p =.024. The categorical measure of willingness to recommend was recoded, collapsing across "don't know" and "no", to form a binary measure indicating whether participants indicated affirmatively that they would recommend the experience. Mirroring the scale measure of recommendation likelihood, less participants in the expectations condition (78.0%) were willing to recommend the movie compared to participants in the control condition (93.8%), χ^2 = 4.96, p = .026.

Satisfaction. The two scale measures of satisfaction were strongly correlated, r = .794, p < .001, and combined to form a single measure. Replicating the recommendation measures, participants in the expectation condition reported less satisfaction with their movie experience than those in the goal condition (M_{goal} = 6.52, SD = 1.05 vs. $M_{expectation}$ = 5.97, SD = 1.54; F (1, 96) = 4.23, p =.042).

This study replicates the effect of expectations on recommendation likelihood in the domain of movie watching. Participants considering their expectations for the movie were less willing to recommend the movie than participants considering their goals. The measures of satisfaction mirrored the recommendation likelihood results.

SUPPLEMENTAL STUDY 3: AMAZON.COM DATA

In studies 5 and 6 of the main text, we used reviews of experiential purchases (hotel stays), and provided correlational evidence that the results obtained in the several lab studies and field experiments presented in the paper hold when we analyze archival data. As we discussed in the paper, our data analysis has some limitations that prevent us from claiming any causal relationship between expectations and review ratings (TripAdvisor) or reviewer recommendations (Expedia). Among these limitations, one may argue that whenever people review purchases, that they are more likely to spontaneously think of and write about expectations when the purchase was disappointing and did not live up to expectations. Such an explanation would imply reverse causality—that negative reviews prompt consideration of expectations. To investigate this possibility, we leverage the findings of study 4 of the main text, which demonstrates that expectations do not decrease evaluations of material goods.

Thus, we collected review data for material goods from Amazon.com, which is a popular platform both for buying material goods as well as reading reviews for such purchases. We coded expectations in the same way in which we coded expectations for the TripAdvisor and Expedia review analyses. If the results of studies 5 and 6 reflect the fact that people are more likely to spontaneously think about and write about expectations when purchases are bad (reverse-causality), then we should expect the results of Amazon.com reviews to replicate the results of the Trip Advisor and Expedia reviews. However, if we have suggested, the earlier results are because, for experiential purchases thinking about expectations reduces one's willingness to recommend the purchase to others, expectations should not decrease star ratings for material purchases on Amazon.com.

This analysis uses a dataset of 24,963 reviews of 4,592 items for the category "All Electronics" obtained from Amazon (McAuley at al. 2015, Hu and McAuley 2016). Using this dataset, we proceed to estimate a specification similar to the one presented in Study 5:

$$Stars_{ijt} = \beta\ Expectation_{ijt} + X_{ijt}\gamma + \alpha_j + \tau_t + \epsilon_{ijt} \qquad (6)$$

where $Stars_{ijt}$ is the star-rating of review i of item j written in year-month t. $Expectation_{ijt}$, whose coefficients is of interest, is an indicator of whether the review i of item j written in year-month t is identified as describing expectations. As usual, we include product and year-month fixed effect, α_j and τ_t, and review environment controls, X_{ijt}.

We report the results of this specification in Supplemental Table 1. In column 1 we present the results without controls and in column 2 we add the review environment controls. In both cases, the coefficient of interest is close to zero and not statistically significant.

Supplemental Table 1: Expectations do not affect star-ratings of non-experiential purchases

	(1)	(2)
Expectations	−0.036	−0.016
	(0.042)	(0.044)
Avg. Stars$_{t-1}$		−0.392***
		(0.040)
log Reviews$_{t-1}$		−0.313***
		(0.031)
N	24963	20371
R^2 within	0.021	0.040

Note: The dependent variable is the star-rating of review i of item j at time t. Cluster-robust standard errors at the individual product level are shown in parentheses.

Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

These results lend further credibility to our interpretation of studies 5 and 6. If worse purchases naturally prompted the consideration and expression of expectations, expectations should have reduced the ratings for the material goods on Amazon.com. Instead, the results are in line with the results of Study 4 and suggest that expectations do not significantly reduce one's propensity to recommend anything, and that this effect is specific to experiential purchases. Moreover, these results suggest that the findings of studies 5 and 6 cannot be explained away as a function of the naivety of our coding system.

STUDY 1 BASELINE MEASURES

1. Have you eaten at this restaurant before?   (0 = No, 1 = Yes)

2. What is your overall impression of the restaurant? (1 star = poor; 5 stars = excellent)

3-6. How would you rate the restaurant according to the following dimensions? (1 star = poor; 5 stars = excellent) 3. Food, 4. Atmosphere, 5. Service, 6. Value

7-9. Please provide your evaluation of the restaurant on the following measures: 7. 1 = Dislike, 7 = Like, 8. 1 = Unfavorable, 7 = Favorable, 9. 1 = Negative, 7 = Positive

10-14. Indicate your agreement or disagreement with the following statements (1 = Strongly disagree, 7 = Strongly agree) 10. I would eat at this restaurant again, 11. I would bring an out-of-town guest to this restaurant, 12. I would recommend this restaurant to someone who seeks my advice, 13. I (sic) say positive things about this restaurant to other people, 14. I would recommend this restaurant to others.

Notes:

Items 2 through 6 were transformed to a 7-point scale (y = x*7/5)

Average quality rating = average of items 2 through 11, Cronbach alpha = .90

Average initial recommendation = average of items 12 through 14, Cronbach alpha = .91

STUDY 3 WRITTEN RESPONSE CODING

Participants written responses about their restaurant visit were coded by participants on Mechanical Turk. An instructional manipulation question was asked at the very beginning of the survey. This question asked participants to indicate that they were taking the time to read

instructions by choosing the following day's date from a list of dates provided. Next, we explained that we were interested in whether statements about a restaurant visit focused on a person's personal experience and emotional reaction or not. Specifically, participants read:

> A description based on someone's experience and emotional reactions may include things like how much fun they had, how delicious/tasty they thought the food was, etc. In contrast, restaurant visits that do NOT focus on the experience and emotional reactions will focus on aspects of the restaurant itself. They will not include how people feel about aspects of the situation and will instead describe things like the aesthetics, good service, high quality food, etc. Such statements are more likely to describe what "it was like" and seem objective in nature.

Participants completed a comprehension check on these instructions. Then, participants were fed twenty statements written by participants (randomized). For each statement, the participant made a binary choice as to whether the statement reflected a personal experience/emotional reaction or not.

Four hundred and thirteen Mturkers completed the survey. 413 people coded the responses. Thirty-one participants failed the instructional manipulation check and were excluded from all analyses. The randomization of statements and the exclusion of people who failed the instructional manipulation check resulted in each statement being rated by between 9 and 33 participants. Responses were coded as focused on one's personal experience if at least half of the respondents indicated that the statement referred to one's personal experience using a binary measure.

SUPPLEMENTAL ANALYSIS OF TRIPADVISOR DATA

Exploiting the year-month of stay

A possible concern with the specification described in Equation 1 (Study 5) is that the time at which the review is written is correlated with changes in hotel quality. For example, if reviews that do not describe expectations are more likely to be written following hotel quality improvement, our results would be biased. To deal with this concern, we estimate the following specification:

$$Stars_{ijt} = \beta\ Expectation_{ijt} + X_{ijt}\gamma + \alpha_j \times Year - Month\ Stayed_{ijt} + \tau_t + \epsilon_{ijt}, \quad (4)$$

where everything is as in Equation 1, except that now we substitute the hotel fixed effect α_j with the hotel – year-month stayed fixed effect. Therefore, the impact of expectations is identified by changes in ratings of reviews of consumers that stayed at the *same* hotel in the *same* year-month. This identification greatly reduces the possibility that reviewers are exposed to different levels of quality. The results reported in Table 5 mimic those obtained using only hotel fixed effects, suggesting that our results are unlikely to be driven by unobserved differences in hotel quality.

SUPPLEMENTAL ANALYSIS OF EXPEDIA DATA

From evaluations to recommendations

In Study 5, we used TripAdvisor ratings and reviews to show that expectations negatively impact ratings. However, one could argue that giving a low rating and not recommending a hotel are not necessarily correlated actions. Here, we exploit the Expedia dataset (Study 6) to show that ratings and recommendations are indeed positively correlated, effectively establishing a link between star-ratings and recommendations.

We test the relationship between ratings and recommendation using the following specification:

$$Pct\ Recommendations_i = \beta\ Stars_i + X_i\ \gamma + \epsilon_i, \qquad (5)$$

where the dependent variable is the percentage of reviewers recommending hotel i as indicated by Expedia. $Stars_i$, whose coefficient is of interest, is the average star-rating of hotel i, and X_i is a vector of controls in which we include a hotel city dummy to account for differences in the percentage of recommendations due to the hotel location that might affect ratings. We present the estimates of Equation 5 in Table 6. In column 1 we present the results without controls. The coefficient of interest is positive and statistically significant, confirming our hypothesis that there is a positive relationship between hotel star-ratings and the probability of recommending the hotel. In column 2, we include the hotel city dummy and obtain a very similar estimate.

Table 5: The effect of expectations on review star-ratings: hotel-month of stay fixed effects

	(1)	(2)	(3)
Expectations	−0.223***	−0.172***	−0.172***
	(0.003)	(0.003)	(0.003)
Affect		0.067***	0.066***
		(0.000)	(0.000)
Avg. Stars$_{t-1}$			−2.344***
			(0.055)
log Reviews$_{t-1}$			0.113***
			(0.020)
N	1766525	1766525	1764954
R^2 within	0.0039	0.072	0.081

Note: The dependent variable is the star-rating of review i of hotel j written at time t. Cluster-robust standard errors at the individual hotel-month level are shown in parentheses.
Significance levels: * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table 6: The relationship between recommendations and expectations

	(1)	(2)
Avg. Stars	20.842***	20.890***
	(0.066)	(0.070)
City Dummy	No	Yes
N	39672	38053
R^2	0.83	0.85

Note: The dependent variable is the percentage of users recommending hotel i. Robust standard errors are shown in parentheses.
Significance levels: * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

WEB APPENDIX REFERENCES

He, Ruining, and Julian McAuley. "Ups and downs: Modeling the visual evolution of fashion trends with one-class collaborative filtering." In *Proceedings of the 25th international conference on world wide web*, pp. 507-517. International World Wide Web Conferences Steering Committee, 2016.

McAuley, Julian, Christopher Targett, Qinfeng Shi, and Anton Van Den Hengel. "Image-based recommendations on styles and substitutes." In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 43-52. ACM, 2015.