Experiences Require Appraisal to Overcome Expectations

Chengyao Sun
Yale University

Jing Wang
Soochow University

Nathan Novemsky
Yale University

This version: April 2025

*** Please contact authors for updated version before circulating or citing.***

Author Note

Correspondence concerning this article should be addressed to Chengyao Sun (chengyao.sun@yale.edu), Yale School of Management, 165 Whitney Ave, New Haven, CT 06511, United States.

ABSTRACT

Beliefs and attitudes about products and brands are presumed to be influenced by experiences with relevant products. In this research, we examine whether and under which circumstances this presumption is correct. In a series of laboratory and field experiments, we show that when belief and experience diverge, experience induces belief updating only when people are nudged to appraise the experience at the time of consumption. Contrary to lay beliefs, surprisingly good and surprisingly bad product experiences have no reliable effect on beliefs and choices when there is no prompt to appraise the experience while it is happening. When there is such a prompt, beliefs and choices shift in the direction consistent with the surprising experience both immediately and several days later. We suggest these results arise because effortful propositional thinking is required to change explicit beliefs (Associative–Propositional Evaluation Model, Gawronski and Bodenhausen, 2006). Our studies suggest that in many experiences, consumers do not expend the effort to articulate their momentary evaluations and therefore, do not update their prior beliefs.

Keywords: appraisal, hedonic evaluation, learning, belief updating, consumer experience

Subjective evaluations and memories of an experience are determined both by the experience itself—i.e., people's sensory response to a particular stimulus— and by expectations about it—i.e., the perceiver's preexisting beliefs and attitudes. How the mind integrates these bottom-up and top-down sources of information has been a major topic of study within psychology and consumer behavior (Cacioppo *et al.*, 1982; Greenwald and Leavitt, 1984; Mazursky and Jacoby, 1986; Park and Smith, 1989; Park *et al.*, 1989; Buschman and Miller, 2007; Ariely and Zauberman, 2000).

The distinction between a person's preexisting beliefs and experiences is especially apparent in cases where the valence of the expectation and the valence of the in-the-moment experience diverge. For example, the rollout of New Coke is infamous because consumers' negative expectations prevented them from appreciating the drink (that was overwhelmingly preferred to classic Coke in blind taste tests; Silver, 2017; Benjamin, 2015). The impact of preexisting beliefs on consumers' evaluations of experiences is widely documented (Bowen *et al.*, 1992; Goldstein *et al.*, 2008; Lee *et al.*, 2006; Nevid, 1981; Plassman *et al.*, 2008; Seymour and McClure, 2008; Wardle and Solomons, 1994). Alongside much evidence for beliefs' overwhelming effects on evaluations, it is obvious that some learning or belief updating occurs and sometimes experiences do shape and shift beliefs: a much hyped movie may disappoint, lowering expectations about the sequel, and a questionable restaurant may deliver a great meal causing patrons to recommend the restaurant to others.

The present research focuses on the role of appraisal prompts at the time of consumption (i.e. asking individuals to reflect on their evaluation of an experience while it is occurring) in moderating the effects of bottom-up influences. Building on the Associative–Propositional Evaluation Model (Gawronski and Bodenhausen, 2006) and research on real-time evaluations (Novemsky and Ratner, 2003), we hypothesize that when expectations and experiences conflict, consumers are often unlikely to *deliberately* evaluate the experience and update their beliefs unless an appraisal prompt is present. In other words, we see experiences influencing beliefs as only occurring under limited circumstances where consumers turn their momentary reactions into explicit notions about the quality of the stimulus.

Next, we review research on the integration of top-down and bottom-up processes and develop our key hypotheses. We then present five experiments that test these hypotheses and some of their implications. We conclude by discussing the theoretical and practical implications of our findings.

THEORETICAL BACKGROUND

The interplay between top-down and bottom-up processes affects attitudes and behaviors in many contexts (e.g., Cacioppo *et al.*, 1982; Greenwald and Leavitt, 1984; Mazursky and Jacoby, 1986; Park and Smith, 1989; Park *et al.*, 1989; Buschman and Miller, 2007; Ariely and Zauberman, 2000; Varelaa *et al.*, 2010; Simonson and Drolet, 2004; Shiv *et al.*, 2005). Often, expectations dominate evaluations of experiences. This is common with brand information (Nevid, 1981), categories like "low-fat" and "full-fat" (Bowen *et al.*, 1992; Wardle and Solomons, 1994), and price (Goldstein *et al.*, 2008; Plassman *et al.*, 2008; Seymour and McClure, 2008;). It occurs in domains ranging from visual perception (Biederman, 1972; Palmer, 1975), to assessment of one's own abilities (Darley and Gross, 1983; Jones *et al.*, 1968), to memories of events (Cohen, 1981; Stangor and McMillan, 1992; Meyvis *et al.*, 2010).

The vast literature highlighting the dominance of expectations suggests that they persist over time and despite conflicting experiences. But common sense, as well as some empirical evidence, suggests that people do sometimes update disconfirmed beliefs (Bolton and Drew, 1991; Cadotte *et al.*, 1987; Oliver, 1980; 1999; Oliver and DeSarbo, 1988; Tse and Wilton, 1988) about products (Oliver, 1977; Swan, 1977), vacation spots (Raghunathan and Irwin, 2001; Swan and Trawick, 1981), and even health choices (Oliver, 1980).

In the present research, we identify some of the conditions that make people more likely to update their beliefs based on direct experience. One apparent difference between studies that find an effect of experiences on belief updating versus those that do not is that studies finding the "bottom-up" effect, by design, have asked participants to rate or appraise their experiences at the time of consumption. For example, Oliver and Burke's multi-stage study revealed that participants updated their beliefs in line with

the quality of the dining experience by having participants record evaluations and describe their thoughts and opinions about the dining experience as it occurred using an audio recorder (Oliver and Burke, 1999; also see Olson and Dover, 1979). Building upon these findings, we propose that being prompted to appraise the experience during consumption facilitates belief updating following a disconfirming experience. This is consistent with research showing that learning from experience is not always spontaneous: to learn from an experience, people need to deliberately attend to it (Hoch and Deighton, 1989), and yet they often do not (Brown and Ryan, 2003). In general, people do not invest cognitive resources unless they feel like they should or are prompted (see Hoch and Ha, 1986; Hoch and Deighton, 1989; Meyvis *et al.*, 2010; Morewedge *et al.*, 2005; Hoch and Deighton, 1989; Weber and Crocker, 1983).

The idea that people do not always appraise their experiences may seem at odds with considerable evidence demonstrating that people can have automatic hedonic reactions to stimuli, and these appraisals can be generated subconsciously (Bechara et al., 1997; Bargh et al., 1996, Duckworth et al., 2002; Fazio et al., 1986). For example, research in the field of social cognition demonstrates that people exposed to a simple stimulus, such as a single word, spontaneously evaluate its valence. This automatic attitude activation is often demonstrated with facilitated word/non-word categorizations of target words following visual or word primes of the same valence in the absence of any explicit evaluation task.

We propose that these discrepant findings can be integrated using the Associative–Propositional Evaluation (or, APE) model (Gawronski and Bodenhausen, 2006). The core notion of the APE model is that implicit and explicit evaluations reflect distinct mental processes—implicit evaluations are based on associative processes, while explicit evaluations reflect propositional processes. The automatic emotional responses represent associative processes, while the explicit learning (or failure thereof) likely represent propositional processes (Fazio and Olsen, 2003; Florack et al., 2001; Wilson et al., 2000). Therefore, although people may have emotional reactions to experiences that challenge their beliefs, they may fail to update those beliefs. To do this updating, the APE model stipulates that consumers need to apply propositional thinking to their spontaneous emotional reactions. Yet this necessary level of propositional

thinking requires effortful cognitive elaboration which may not always or even generally occur spontaneously. In this case, an external nudge or what we call an *appraisal prompt* might be needed to capture an otherwise fleeting affective response and facilitate belief updating. Therefore, we propose our main hypothesis:

> H1: A prompt to evaluate a disconfirming experience at the time of consumption causes belief updating that would not occur otherwise.

One crucial element in our prediction is that appraisal prompts operate by cueing people to make the effort to engage propositional thinking that is not otherwise engaged. Our effects are not driven by a shift in attention to the experience per se, but rather by nudging consumers to think about their own affective responses to the experience. Indeed, some previous research suggests that people may spontaneously attend to experiences (Morewedge *et al*., 2010). But we believe it is their ephemeral positive or negative responses to the experience that consumers may not spontaneously articulate and therefore may not get factored into their beliefs. Prompts to evaluate an experience induce propositional thinking about these reactions which in turn facilitates belief updating. Following this logic, merely asking people to focus on other aspects of the experience such as prompting people to evaluate specific attributes of the stimulus may not produce the same effects because this will not cause people to articulate their affective responses to the experience.

Our theory implies that without a prompt and the propositional thinking that follows, memory for the evaluation of a particular experience will likely fade. Therefore, later explicit judgments of that specific experience (e.g., how good was that piece of fat-free cheese you ate in the lab last week) will be largely driven by preexisting beliefs about the category that contains that experience because the quality of that specific experience was never encoded via propositional thinking into explicit knowledge.

One key mechanism of our theory is that our prompts induce people to engage in a propositional thinking process about their affective responses to an experience. If true, any manipulation that causes

people to deliberate on their affective responses to the experience, with or without overt reporting of the evaluation, should cause belief updating. Therefore, we propose the following hypothesis:

H2: Appraisal prompts without explicit reporting of experience evaluations causes belief updating that would not occur otherwise.

For our central predictions outlined as H1 and H2, we do not specify which particular belief gets updated. We think consumers are likely to update the belief that happens to be the most accessible at the moment of appraisal. Since accessibility is influenced by a variety of factors including the particular choice contexts (e.g., the specific wording of appraisal prompts) and the consumers' chronically accessible thoughts (Biehal and Chakravarti, 1983), we should see a variety of beliefs that could get updated, including those about a specific product, brand, or the product category.

However, a single experience might not always lead to belief updating even when attention is drawn to one's affective reactions to the experience. For instance, a more strongly held belief, perhaps because it is about a broader category, could be too entrenched to change just from a single surprising experience even in the presence of an appraisal prompt. Imagine you just saw an unattractive designer product from a famous and well-liked brand. Articulating affective responses while you are viewing the item might not be sufficient to change your opinions about the brand but should shift your evaluation of the specific item. Hence, we propose breadth of category as a moderator such that an appraisal prompt is more likely to facilitate belief updating about a narrow than a broad category. More formally, we propose the following hypothesis:

H3: When a disconfirming experience is an exemplar of both a narrow and a broad category, a prompt to evaluate the experience during consumption increases belief updating more for the narrow than the broad category.

Although we argue that propositional thinking about affective responses to an experience may not be the default, we do believe that some experiences will induce propositional thinking without any external prompting. According to our theorizing, any factor that motivates people to deliberate on their affective responses during an experience and capture these otherwise fleeting feelings through propositional thinking would facilitate belief updating. One such factor could be when consumers are actively considering a pending purchase. For example, when consumers are actively considering an imminent purchase, one question they might be asking themselves during a trial experience is "how much do I like this?" Notice, the knowledge that they can purchase the item being experienced is not the same as active consideration of whether the item is worthy of purchase. Hence, reminding consumers about a pending purchase opportunity should produce similar effects as our appraisal prompt by motivating them to form propositional thoughts around the focal item and in turn lead to belief updating without a nudge specifically to appraise the experience.

H4: Reminding consumers about an imminent purchase opportunity during an experience increases belief updating.

OVERVIEW OF STUDIES

We report five studies that test our hypotheses. We begin our investigation by testing our central prediction about the role of appraisal prompts at the time of consumption in updating subsequent beliefs. Study 1 examines whether people only update their beliefs following a disconfirming experience when prompted to articulate their feelings during the experience (H1). Study 2 extends Study 1 by examining whether the existence of an accurate memory of a specific surprising experience depends on an appraisal prompt at the time of consumption. Study 3 employs an appraisal prompt without any overt reporting of the appraisal to test whether reporting of affective responses during an experience is necessary for belief updating (H2). Study 4 tests a boundary condition of our effects by examining whether a prompt to

evaluate an experience is more likely to result in belief updating for a narrow rather than a broad category (H3). Finally, Study 5 examines when people might update their beliefs without an appraisal prompt. We test whether reminding people of a pending purchase opportunity makes an appraisal prompt no longer necessary for belief updating (H4). Our study materials, data, and code are available at: https://osf.io/fk5hd/?view_only=edc69d836e2a41c8800db3ca4beea6e4.

STUDY 1: EFFECT OF APPRAISAL PROMPT ON BELIEF UPDATING

To test hypothesis 1, we examine the change in people's beliefs following an experience that diverges from expectations with or without appraisal prompts. Specifically, we asked participants to sample a fat-free cheese that tasted better than expected. We varied whether or not participants were prompted to appraise their experience during consumption. One group was prompted to rate their liking of the cheese as they were eating (Appraisal Prompt condition), while the control group did not. We then examined people's beliefs about fat-free cheese both immediately following the experience and several days later.

Our theorizing relies on the idea that appraisal prompts will generate propositional thoughts around one's emotional reactions to an experience. Therefore, a focus of attention on the experience per se should not drive belief updating. It is important to compare the appraisal prompt condition to one in which participants are focused on the experience of eating the cheese but not on their emotional reactions to it. To this end, we asked a third group of participants to rate the physical traits of the cheese during consumption (Physical Traits condition).

Method

Respondents were one hundred and forty-three undergraduate students at a major West Coast university who participated in this and a number of other studies for partial course credit. The study used

a three cell (Appraisal Prompt vs. Physical Traits vs. Control) between-participant design. Participants were run in groups, with approximately 20 -25 people per session. Participants were told that the company sponsoring this study had a new type of cheese that was developed as a substitute for cheese containing fat. The company was said to be pre-testing certain aspects of the product with small, targeted populations of consumers, of which students were one group. Participants were then asked a few questions regarding their cheese consumption, including how much they like fat-free cheese (i.e., their pre-existing beliefs toward the category) on a 9-point scale (1 = strongly dislike, 5 = neutral, 9 = strongly like). They also were screened for medical or other factors that would preclude them from eating cheese containing fat. Three participants reported an aversion to sampling cheese containing fat and were dropped from the study. Another participant did not indicate their pre-existing beliefs about fat-free cheese and was thus also excluded, leaving us with a final sample of 139 respondents.

After answering these pre-trial questions, each student then received a single slice of individually wrapped American cheese that contained fat (i.e., to ensure that the taste would be better than expected) and was told the cheese was a new type of fat-free cheese.

Participants in the Appraisal Prompt condition responded to two questions assessing their liking of the experience: “How much do you like the fat-free cheese you are eating?” (1 = not at all, 5 =somewhat, 9 = very much) and “How does this cheese compare to your expectations about fat-free cheese? (1 = much worse than expected, 5 = about what you expected, 9 = much better than expected). Participants in the Physical Traits condition responded to two questions assessing their evaluation of the texture and color of the cheese: “Do you think the texture of this cheese is similar to cheeses you have had in the past?” “Do you think the color of this cheese is similar to cheeses you have had in the past?” (1 = not at all similar, 5 = somewhat similar, 9 = very similar). These dimensions were chosen because they are closely related to evaluations in the sense that color and texture are two dimensions on which fat-free cheese can be different from other cheese and these differences could account for the lower expectations for fat-free cheese compared to other cheese. Participants in the Control condition responded to two questions assessing their evaluation of their classes and study habits: “How much are you learning in your

classes this semester?" (1 = much less than expected, 5 = about what you expected, 9 = much more than expected) and "How much time do you spend studying compared to your classmates?" (1 = much less than others, 5 = about the same as others, 9 = much more than others). This last condition was meant to mirror a real consumption situation where consumers are often thinking about other things while consuming a product.

Note that in the Appraisal Prompt condition, in addition to liking, we also asked participants to compare their real-time experience with their prior expectation. This question is designed to parallel the comparative evaluations used in the Physical Traits and Control condition. We show in our later studies that explicit mention of expectations is not necessary for our effects to hold.

Following the sampling experience, each participant completed a brief questionnaire measuring 1) their liking and 2) their purchase intentions for products in several unrelated food categories (e.g., yogurt, frozen waffles, peanut butter). Embedded within this questionnaire were two items on which participants reported how much they like fat-free cheese in general and how likely they would be to purchase fat-free cheese in the next two weeks. We asked purchase intentions at this point to try to avoid any demand to be consistent with ratings given during the experience (i.e., there are many reasons purchase intentions could diverge from liking of a product).

Respondents in all conditions were then sent a brief follow-up email either 3 or 10 days later. To minimize attrition, we only asked participants to indicate how much they like fat free cheese in general on a 9-point scale (1 = strongly dislike, 9 = strongly like). All but one participant responded to this delayed assessment. The length of delay did not impact any of the results reported here, so we collapse across the length of delay in the analyses that follow[1].

[1] The separate means for each condition between a delay of 3 days and 10 days are as follows: $M_{\text{Appraisal Prompt 3-day}}$= 5.52/ $M_{\text{Appraisal Prompt 10-day}}$= 5.58; $M_{\text{physical traits 3-day}}$ = 4.78/ $M_{\text{physical traits 10-day}}$ = 4.72; $M_{\text{control 3-day}}$ = (no data)/$M_{\text{control 10-day}}$ = 4.63

Results

Manipulation check. Participants who rated their liking of the fat-free cheese sample during consumption indicated that liking of the cheese exceeded pre-existing beliefs about fat-free cheese (real-time liking M = 6.39 versus pre-trial ratings of fat-free cheese M = 4.73, t(55) = 4.97, $p < .01$). They also directly indicated that the sample tasted better than what they expected when asked specifically to compare the experience to expectations (M = 6.55, t(55) = 6.62, $p < .01$ against the mid-point 5 labeled "about what you expected"). As expected from random assignment, participants' pre-trial beliefs about fat free cheese did not differ across conditions ($M_{appraisal\text{-}pre\text{-}trial}$ = 4.73; $M_{physical\ traits\text{-}pre\text{-}trial}$ = 4.82; $M_{control\text{-}pre\text{-}trial}$ = 4.52, F(2, 136) = .23, $p = .80$).

Belief about fat-free cheese after a delay. An omnibus ANOVA revealed significant differences between conditions in our key measure—people's beliefs about fat-free cheese after the delay (F(2, 136) = 4.60, $p = .01$). A planned contrast confirmed that beliefs about fat-free cheese were more positive for those who appraised their experience ($M_{appraisal\text{-}delay}$ = 5.55) than for the other two groups combined ($M_{physical\ traits\text{-}delay}$ = 4.75, $M_{control\text{-}delay}$ = 4.63, t(136) = 3.01, $p < .01$). There was no reliable difference between the Physical Traits and Control groups (t(136) = .32, $p = .95$).

Comparing beliefs about fat-free cheese prior to the tasting experience versus several days after, we find no significant change in beliefs about fat-free cheese for those who made ratings unrelated to the cheese during consumption ($M_{control\text{-}pre\text{-}trial}$ = 4.52, $M_{control\text{-}delay}$ = 4.63, t(26) = .83, $p = .42$). The same comparison among those who rated the color and texture of the cheese also shows no reliable change in beliefs about fat-free cheese ($M_{physical\ traits\text{-}pre\text{-}trial}$ = 4.82 vs. $M_{physical\ traits\text{-}delay}$ = 4.75, t(55) = .56, $p = .58$). These results suggest that a single surprising experience on its own may be insufficient to induce belief updating (See Figure 1).

However, participants who were prompted to rate their liking of the cheese at the time of consumption show a very different pattern of results, revealing significantly improved evaluations of fat-free cheese relative to their preexisting beliefs ($M_{appraisal\text{-}pre\text{-}trial}$ = 4.73 vs. $M_{appraisal\text{-}delay}$ = 5.55, t(55) = 3.96, p

< .01). The updating shown in the Appraisal Prompt condition suggests that a single experience with one slice of cheese is sufficiently informative to warrant updating attitudes toward fat-free cheese, but this updating only occurs when people articulate their feelings during the experience. Collectively, these results provide support for H1.

Immediate category liking and purchase intentions. We did not make specific predictions about effects of appraisal prompts on immediate post-trial belief updates. We were not sure if feelings from the experience might still be accessible in memory and incorporated into the immediate post experience belief measures. In principle, we believe if the prompt causes people to think about the experience they just had and their feelings about that experience are still accessible, then we will see updating. If either of these are missing, then there should be no updating. To test whether immediate evaluations of fat-free cheese showed the same pattern as delayed evaluations, we examined liking and purchase intentions reported after sampling the fat-free cheese in the initial experimental session. An omnibus ANOVA revealed significant differences between groups in purchase intentions ($F(2, 136) = 4.01$, $p = .02$). A planned contrast confirmed that intentions to purchase fat-free cheese even just a few minutes post-consumption were higher for those who made a real-time appraisal of the tasting experience ($M_{appraisal\text{-}prompt} = 4.46$) than the other two groups combined ($M_{physical\text{-}traits} = 3.23$, $M_{control} = 3.67$, $t(136) = 2.46$, $p = .02$). There was no reliable difference between the Physical Traits and Control groups ($t(136) = .80$, $p = .70$).

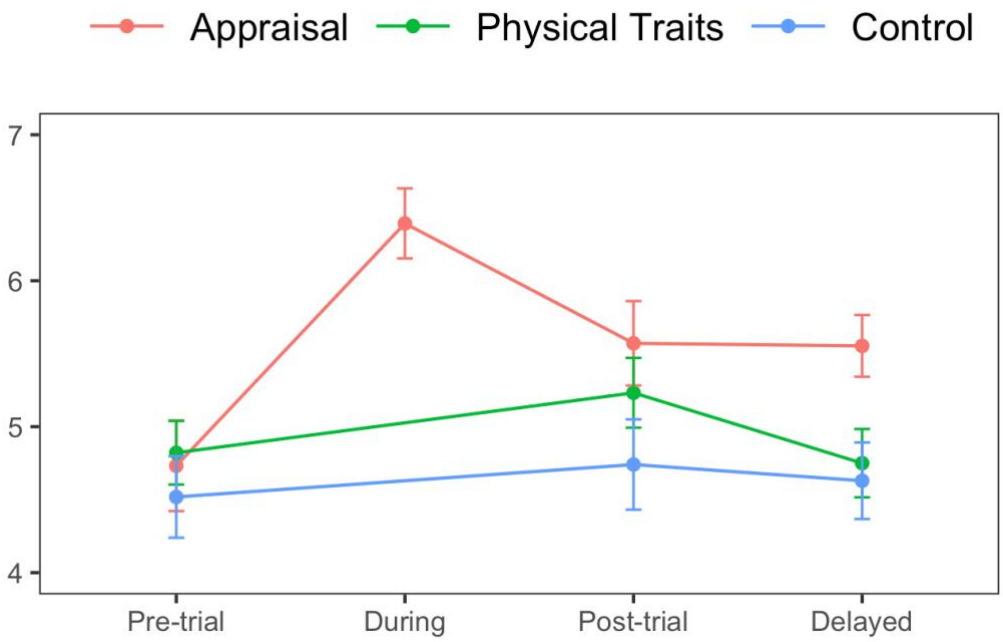
Liking of fat free cheese was reliably higher immediately following the tasting experience in the Appraisal Prompt ($t(55) = 3.42$, $p < .01$) condition compared to pre-trial ratings. There was no reliable increase in the control condition ($t(26) = 1.44$, $p = .16$), suggesting any affect felt during the tasting experience faded right away in the control condition. Liking for the fat-free cheese increased immediately in the Physical Traits condition ($t(55) = 3.38$, $p < .01$), but it faded away in a few days ($M_{post\text{-}trial} = 5.23$, $M_{delayed} = 4.75$, $t(55) = 2.58$, $p = .01$). Since participants gave exceptionally high ratings for the physical traits ($M_{texture} = 7.53$, $M_{color} = 8.32$, on scales of 1-9)—presumably because the sample was very similar to the non-fat-free cheese people usually had—it was possible that they made positive inferences from their high ratings when asked about liking shortly after. Whatever was picked up in these liking ratings affected

neither their long-term beliefs about fat-free cheese nor their contemporaneous purchase intentions, as they indicated the lowest purchase intent among the three conditions.

Table 1 Study 1 Means

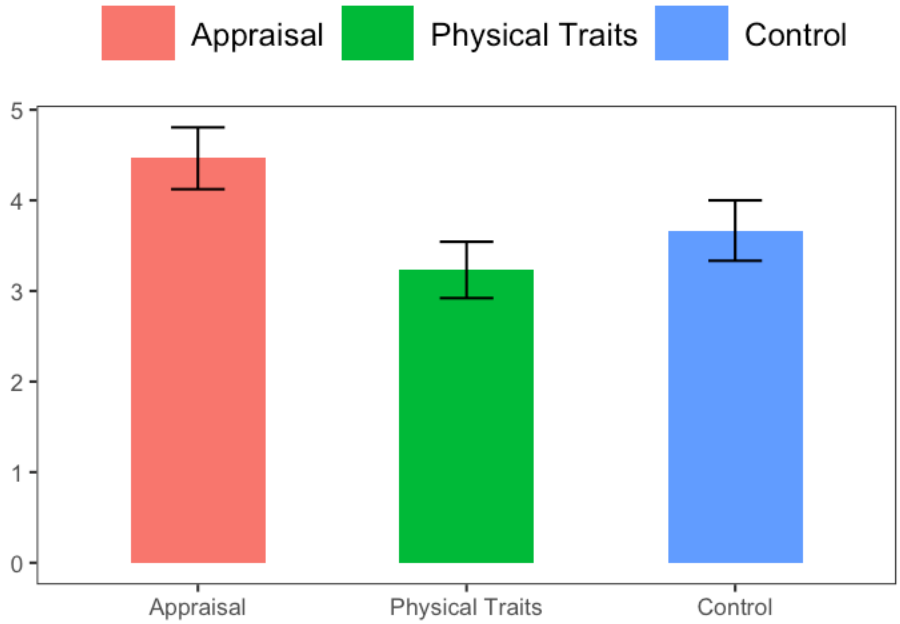
Rated Liking	$M_{Appraisal\ Prompt}$	$M_{Physical\text{-}Traits}$	$M_{Control}$
Pre-trial	4.73	4.82	4.52
During	6.39	—	—
Post-trial	5.57	5.23	4.74
Delayed	5.55	4.75	4.63
Post-Trial Purchase Intentions	4.46	3.23	3.67

Figure 1 Study 1 Effect of Appraisal Prompt on Belief Updating Over Time



Note. Error bars indicate +/- one standard error.

Figure 2. Study 1 Post-Trial Purchase Intentions



Note. Error bars indicate +/- one standard error.

Discussion

The results of Study 1 show that a disconfirming experience is more likely to facilitate learning and belief updating both immediately and following a delay when individuals are prompted to make a real-time appraisal during the consumption experience. In the absence of such an appraisal prompt, beliefs showed no change even a few minutes after the experience. In the presence of such a prompt, we found significant updating of beliefs and a significant change in purchase intentions toward the focal product category (fat-free cheese). Consistent with our theorizing, evaluating other aspects of the experience (i.e., color and texture) are not sufficient to bring about long-term changes in beliefs. The differences in the pattern of results between the Appraisal Prompt and Physical Traits conditions highlight that our effects are not driven by a shift in focus of attention to the experience per se, but rather it is the ephemeral affective response to the experience that our prompts cause participants to capture in propositional thinking that accounts for our results. When the prompt is not about one's own affective reactions to the

experience, but merely about the experience, those fleeting feelings may not get captured by the propositional thinking that is necessary for belief updating.

Findings from Study 1 raise an interesting issue regarding people's intuitions or lay theories about when learning will result from experience. Do people intuit the effect of appraisal prompts on belief updating or do they believe a relevant experience alone is sufficient to shift their beliefs? We suspect both marketers and consumers overestimate the impact of a single improved user experience on belief updating. Managers do not always understand the key drivers of their customers' satisfaction (Hult *et al.*, 2017), and consumers' lay intuitions about drivers of satisfaction have been shown to diverge from reality in various domains (e.g., Schkade and Kahneman, 1998).

To test this idea, we ran a follow-up study with four hundred twenty-two individuals recruited through a survey website. The experiment consisted of a 2 (product quality: typical vs. good fat-free cheese) by 2 (appraisal prompt: present vs. absent) between-participant design. Participants read about individuals who sampled either a surprisingly good-tasting or typical fat-free cheese and were told either that the individuals were or were not asked to "make a liking rating of this cheese while eating it." Participants were then asked to indicate how the individuals' liking and purchase intent for fat-free cheese would change "from before sampling the fat-free cheese to after sampling the fat-free cheese." Our results revealed only a main effect of product quality on both belief updating and purchase intent, indicating that people think that eating a good tasting fat-free cheese will impact beliefs and purchase intent more than eating a typical fat-free cheese. More importantly, participants did not think appraisal prompts would make any difference in learning.

STUDY 2: EFFECT OF APPRAISAL PROMPTS ON MEMORIES OF A PARTICULAR EXPERIENCE

Our theory suggests if there is no appraisal prompt, then the affect from a surprising experience will decay without being converted into propositional thinking and without being stored as part of one's

explicit beliefs. In the absence of such memory, later recollection of that specific experience should be largely driven by preexisting beliefs about that type of experience. In the presence of an appraisal prompt, a person's otherwise fleeting positive or negative reactions to the experience will more likely be incorporated into explicit beliefs and in turn influence memories of that experience. In Study 2, we examine memories of an experience under conditions where appraisal is prompted (vs. not) to test whether memories of experiences are bottom-up in the presence of an appraisal prompt and top-down in the absence of that prompt. In addition, we wanted to test another domain of experiences to extend our findings beyond food consumption. Lastly, Study 2 extends our investigation to a surprisingly bad (rather than surprisingly good) experience. Ample research has shown that negative experiences loom larger than positive ones in influencing customer satisfaction and are likely to draw more spontaneous attention (see Anderson and Sullivan, 1993 for a review). While we found that spontaneous belief updating was unlikely in the case of a surprisingly good experience without appraisal prompts in Study 1, we wanted to test if a surprisingly bad experience might be more likely to induce spontaneous bottom-up memories.

Method

Participants were students at a major West Coast university taking the study for partial course credit. Participants were run in groups of approximately 20 to 25 people per session. Participants first evaluated how much they liked or disliked each of seven movies, including "The Matrix", on a scale of -5 (extremely dislike) to 5 (extremely like). We used this preliminary question to gauge their pre-existing beliefs about "The Matrix". Participants might leave a question unanswered if they did not watch that movie or for other reasons. One hundred seventeen participants left a rating for The Matrix. This study thus focused on these 117 participants who indicated their pre-existing beliefs about "The Matrix".

Participants were shown a short scene from "The Matrix." The movie and scene were chosen based on a pilot study that revealed this particular movie was well-liked (M_{Movie} = 8.83 on an 11-point

scale where 11 indicated extreme liking) among members of the sample cohort. The test scene was one of the least interesting and least-liked scenes in the film, and considered far less enjoyable than the movie as a whole (M_{scene} = 5.6, t(21) = 3.12, p = .005). In this way, we drew a surprisingly bad experience (scene) from what is otherwise a good category (movie).

After viewing the scene which lasted approximately four minutes, half the participants were prompted to evaluate the scene while the other half completed unrelated evaluations. Of those participants evaluating the scene, half were asked to rate the "action clip" they just saw and half were asked to rate the "clip from the film 'The Matrix.'" We varied the language to examine whether indicating the scene was from a well-liked movie might affect evaluations reported for the scene. One week later, all participants were sent an e-mail asking them to evaluate the clip they had seen from memory. We obtained responses from 109 of the 117 participants. Half of those who did not evaluate the clip at the time of viewing were asked to rate the "action clip," while the other half rated the "clip from the film 'The Matrix.'" Those participants who had already rated the clip were asked to rate it again and were given the same label they had been given earlier.

Results and Discussion

All participants rated the clip on four separate scales. They rated how much they like it and how much they would like to see more scenes like it on a scale of -5 to 5 with 0 labeled as the neutral point. They also rated how entertaining they found it and how engaging it was on a scale of 0 to 11. Individual responses to these four questions were summed to produce a single rating of the clip at each point in time (Cronbach's α = .76 for immediate likings, α = .82 for delayed evaluations). No significant difference

emerged between those who saw the label "action clip" and those who saw the label "The Matrix," so we collapsed across this variable in all reported analyses[2].

One week after viewing the scene, those who were not prompted to rate the scene in real-time thought it was significantly better (M = 15.7) than how the scene was rated in real-time (M = 10.5, t(111) = 3.75, $p < .01$). Those who were prompted to rate the scene in real-time showed no significant bias in memory (M = 10.5 in real-time versus M = 9.7 in retrospect, t(54) = 1.04, $p = .30$). Correlational analyses showed that evaluations of the movie collected before viewing the scene reliably correlated with delayed evaluations of the disappointing scene only if participants were not prompted to make real-time evaluation of the scene (r = .53, $p < .01$). For participants who were prompted to make real-time evaluations of the scene, neither real-time evaluation of the scene nor their memory of the scene correlated with their overall beliefs about the movie (r = .05, $p = .73$ and r = -.19, $p = .17$ respectively). Instead, their memory of the scene reliably correlated with their real-time evaluation (r = .40, $p < .01$). These results provided further evidence that without an appraisal prompt, feelings during an experience leave little trace, and subsequent judgments are based on pre-existing beliefs.

STUDY 3: EFFECT OF UNSTATED APPRAISAL ON BELIEF UPDATING

According to our theory, the appraisal prompts operate by inducing participants to engage propositional thinking about their affective responses to an experience. If true, this suggests that any manipulation that causes people to pay attention to their own affective evaluation of the experience, with or without overt reporting of the evaluation, should cause belief updating. Study 3 was designed to test the idea that reporting of affective responses is not needed for belief updating (H2). We compare three

[2]The real-time mean rating for "Matrix" is 11.3 and 9.7 for "action clip" t(57) = .76, p =.45; the delayed mean rating for "Matrix" without real-time appraisal prompt is 14.6 and 16.8 for "action clip" t(52) = 1.29, $p = .20$; the delayed mean rating for "Matrix" with real-time appraisal prompt is 9.9 and 9.4 for "action clip" t(53) = .32, $p = .75$.

conditions: one in which participants are prompted to rate their experience; a new condition in which participants are encouraged to pay attention to how they feel about the experience in real-time but are not asked to report their feelings, and a third condition in which participants are not given any appraisal prompt.

Additionally, in Studies 1 and 2 it is possible that the effects of appraisal prompts were driven in part by the fact that the scales used to assess real-time evaluations in the Appraisal Prompt condition were very similar to the scales used to assess post-consumption beliefs. This could have created potential demand for participants to later report beliefs that were consistent with their real-time evaluation. Since Study 3 has a condition without any reporting during the experience, the results can shed light on this alternative explanation. This study is pre-registered at AsPredicted: https://aspredicted.org/pt8f-jj3k.pdf.

Method

We recruited seven hundred fifty-one participants from Prolific.com, an online survey platform. Because this study used audio and video, four participants who reported that they were unable to use video or audio were excluded from the study, as pre-registered, leaving us with a final sample of 747. Participants were asked to watch a 30-second clip of Jerry Seinfeld's stand-up comedy. The comedian and the clip were chosen based on a pilot study revealing that the clip was less funny than this comedian's stand-up comedy in general (M_{clip} = 3.1, $M_{general}$ = 5.1, on a scale of 1-9 where 9 indicated extreme funniness, $t(97) = 5.1$, $p < .01$).

Participants were randomly assigned to one of three conditions. Participants in the Stated Appraisal condition were instructed, "As you watch the stand-up comedy, please think about how funny it is." Immediately after the clip stopped, they were asked to rate how funny they think the clip was on a scale of 1 (not at all funny) to 9 (very funny). Participants in the Unstated Appraisal condition were also instructed, "As you watch the stand-up comedy, please think about how funny it is," but they were not asked to rate it. Participants in the Control condition was not prompted to think about or rate how funny

the clip was. Instead, they rated how often they participated in surveys that involve videos on scale of 1 (never) to 9 (very frequently).

Next, all participants completed three sets of filler questions about movies, music, and gourmet food. Finally, all participants then rated how much they would enjoy watching Jerry Seinfeld's stand-up comedy on a scale of 1 (not at all) to 9 (very much).

Results and Discussion

A one-way ANOVA showed that participants in the three conditions reported different ratings of how much they would enjoy watching Jerry Seinfeld's stand-up comedy ($F(2, 744) = 13.61$, $p < .01$). A planned contrast revealed that participants indicated less enjoyment of Jerry Seinfeld's stand-up comedy in the Stated Appraisal condition ($M_{stated} = 4.4$) than in the Control condition ($M_{control} = 5.4$, $t(744) = 4.74$, $p < .01$). More importantly, another planned contrast revealed that participants also indicated less enjoyment in the Unstated Appraisal condition ($M_{unstated} = 4.5$) than in the Control condition ($M_{control} = 5.4$, $t(744) = 4.25$, $p < .01$). The rated enjoyment did not differ between the Stated and Unstated Appraisal conditions ($t(744) = .50$, $p = .871$). All analyses are pre-registered.

The results of Study 3 supported H2, and replicated and extended the results of Studies 1 and 2 by demonstrating that appraisal prompts facilitate belief updating without the reporting of those appraisals. Merely thinking about one's feelings about a disconfirming experience causes belief updating. Further, the results of this study showed that participants who received the "unstated" versus "stated" prompt rated their enjoyment of Jerry Seinfeld's comedy similarly, suggesting as we hypothesized, that the key factor driving our effects would be making an appraisal rather than providing a rating to the experimenter. These findings cast doubt on alternative explanations based on response consistency.

STUDY 4: CATEGORY BREADTH AS A MODERATOR

So far, we have shown that real-time appraisals facilitate belief updating in various domains, including food, movies, and comedy. However, some beliefs are more strongly held and must be harder to update than others. A broad and highly familiar category for instance could be among the hardest to update because these beliefs could be based on many experiences and hence too entrenched to get updated from a single surprising experience, even in the presence of an appraisal prompt. Study 4 is designed to test this boundary of our effects, by orthogonally varying the presence versus absence of an appraisal prompt and category breadth. We predict that a single experience will have a greater impact on a narrower category belief compared to a broader one. In addition, this effect should be particularly pronounced among those with strong beliefs about the broad category. This study is pre-registered at AsPredicted: https://aspredicted.org/s754-gn6k.pdf.

Method

This study examined consumer beliefs about the luxury fashion brand Christian Dior and its products. We pre-registered to recruit 480 female participants because research shows that females are more involved in fashion than men and therefore should have more entrenched beliefs about luxury fashion brands (e.g., Soeck and Bailey 2007). We posted our study on Prolific.com and set a prescreening criterion so that only females could see our posting and participate in our study. At the end of study, we asked participants to indicate their gender again. Four hundred seventy-nine participants completed the study. Three indicated that they were males at the end of study and thus were excluded, as pre-registered, leaving us with a final sample of 476 female respondents.

Participants were randomly assigned to one condition of a 2 (appraisal prompt vs. no appraisal prompt) by 2 (narrow vs. broad category) between-participant design. Participants were first shown a picture of a Christian Dior limited-edition bracelet that had been pretested to be particularly unappealing. They were told that the bracelet was a limited edition from Christian Dior. Half of the participants assigned to the appraisal prompt conditions were asked to rate 1) how much they liked the bracelet on an

11-point scale ranging from -5 (I do not like it at all) to 5 (I like it very much) and 2) how much they liked the bracelet compared to Dior's other designs on an 11-point scale ranging from -5 (much less than their other designs) to 5 (much more than their other designs). The other half of the participants assigned to the no appraisal prompt conditions were asked whether they had seen it before. All participants were then given two filler word puzzle games that lasted for at least 6 minutes in total. Participants in the narrow category conditions were then asked to indicate how much they liked Dior's limited-edition bracelets on a 7-point scale ranging from "Dislike a great deal" to "Like a great a deal". Participants assigned to the broad category were asked to indicate how much they liked the fashion brand Dior on the same 7-point scale. We predicted that an appraisal prompt would lead to more belief updating about Dior's limited-edition bracelets than to updating for the brand in general.

Results and Discussion

As pre-registered, we regressed participants' posterior beliefs on the appraisal prompt (0 = absent, 1 = present), the category breadth (0 = narrow, 1 = broad), and their interaction. The regression results showed a significant negative coefficient of the appraisal prompt (b = -.43, SE = .19, p = .02), suggesting that the real-time appraisal prompt facilitated belief updating for the narrow category (i.e., when the category breadth variable was held at 0). The regression results also showed a marginally positive interaction (b = .50, SE = .28, p = .07), suggesting that the effect of the appraisal prompt was smaller for the broad category than for the narrow category. To better understand how the appraisal prompt exerted different effects on belief updating for narrow and broad categories, we conducted two simple effects tests. First, a contrast comparing the narrow category beliefs in the presence and absence of an appraisal prompt showed a significant difference whereby people liked Dior's limited-edition bracelets significantly less after viewing the unattractive piece with an appraisal prompt ($M_{appraisal\text{-}narrow}$ = 3.33) than without an appraisal prompt ($M_{no\ appraisal\text{-}narrow}$ = 3.76, t(346) = -2.03, p = .04). This replicated our previous findings for effects of appraisal prompts. Second and more importantly, we conducted a contrast comparing beliefs

about the broad category with and without an appraisal prompt. We found no significant difference between how much people liked the luxury brand Dior with ($M_{appraisal\text{-}broad}$ = 4.64) or without an appraisal prompt ($M_{no\ appraisal\text{-}broad}$ = 4.57, t(346) = .36, p = .72).

Supporting H3, this study demonstrates that breadth of category can moderate the effects of appraisal prompts on belief updating: viewing a bad design from a luxury brand leads to more updating of beliefs about the design than beliefs about the brand in the presence of appraisal prompts. This is consistent with the idea that strongly held beliefs are harder to update even in the presence of an appraisal prompt. Broad category beliefs could be difficult to update in part because these beliefs are about a high involvement familiar category and hence too entrenched to update from a single surprising experience.

STUDY 5: COMPARING AN APPRAISAL PROMPT TO A PURCHASE PROMPT

Given the results thus far, one might naturally wonder: does belief updating ever occur without an appraisal prompt? In other words, when would people spontaneously generate propositional thoughts without an external nudge to evaluate their experience? From a managerial perspective, it is vital to understand the conditions under which an improved customer experience would translate into elevated brand beliefs. Furthermore, whether belief updating translates to increases in choice share would be a critical outcome for managers. Study 5 is designed to examine these issues.

We propose that reminding people of an imminent purchase opportunity will motivate them to pay attention to their own affective reactions during an experience and to engage in propositional thinking necessary for belief updating without any additional nudge for appraisal (H4). Future purchase opportunities are generally possible for most product experiences, but these opportunities may not be top of mind during an experience. For instance, people may not be actively thinking of whether to return for another visit to an amusement park or a restaurant while in the park or dining at the restaurant, although they generally know that this is possible.

In this study, we provide people with a positive experience that challenges their preexisting

beliefs about a product and they are either reminded of an opportunity to purchase the product before the sampling experience or not. We give participants a real choice to purchase the product to test whether appraisal prompts impact a consequential choice. We predict that when explicitly considering a pending purchase, people are likely to pay attention to their own affective reactions during an experience (i.e., asking themselves how much do I like this coffee?) and generate propositional thoughts about the sampling experience to be incorporated into their existing beliefs. If true, adding an appraisal prompt for these participants should have no further effect.

Method

Study 5 used a 2 (appraisal prompt vs. no appraisal prompt) by 2 (purchase reminder vs. no purchase reminder) between-participant factorial design. We set up Study 5 to mirror a real-world product sampling situation. We chose instant coffee as the focal product. The general consensus among our target population is that instant coffee doesn't taste very good compared to other readily available types of coffee such as freshly brewed coffee. This particular brand of instant coffee[3] was chosen based on a preliminary blind tasting session held with the same target population where it was deemed hardly discernable from freshly brewed coffee. Participants were given a choice between keeping a 100 NTD (New Taiwan Dollar) cashcard or exchanging it for an 80g jar of the instant coffee they just tasted, which retailed at the time for 130 NTD.

A table was set up on the campus of a major Southeast Asian University with a big sign saying "FREE COFFEE". Data was collected over the course of three weeks every day around 11am to 2pm. Sample jars of the coffee were displayed on the table.

Three hundred and two students and staff members participated in this study. Participants were first offered free coffee in an unmarked paper cup. Together with the coffee sample, every participant was

[3] A Japanese brand named AGF was used in this study. This particular product configuration was introduced in the local market about 6 months before the study.

also handed a 100 NTD cashcard redeemable at the campus convenience store or any store of the same chain around the city. About a quarter of the participants (Control condition) were handed a cashcard and drank the coffee. Only after trying the coffee did we ask them whether they would like to keep the cashcard or to use the cashcard to buy an 80g jar of the same instant coffee they just tasted. Another quarter of the participants (Purchase Prompt condition) were handed a cashcard, and were told that they would later be given an option to either keep the cashcard or use the card to buy an 80g jar of the same instant coffee they were about to drink. They then drank the coffee and made a choice between keeping the cashcard or buying the coffee. Another quarter of the participants (Appraisal Prompt condition) were handed the cashcard, drank the coffee and while drinking the coffee were asked to evaluate how much they liked the coffee they were drinking on a scale ranging from 1 (do not like at all) to 10 (like very much). They then made the same choice between keeping the cashcard or buying the coffee. A final quarter of the participants (Purchase +Appraisal Prompt condition) were handed a cashcard and told that they would later be given an option to either keep the cashcard or use the card to buy an 80g jar of the same instant coffee they were about to drink. They then drank the coffee, and while drinking the coffee were asked to evaluate how much they liked the coffee they were drinking on a scale ranging from 1 (do not like at all) to 10 (like very much). They then made the choice between keeping the cashcard and buying the coffee.

Results and Discussion

Compared to the Control condition (N = 76), where 35.5% chose to buy the instant coffee, participants were more likely to choose instant coffee over the cashcard in all three of the other experimental conditions. Specifically, 63.2% of the participants in the Appraisal Prompt Condition (N = 76), 57.3% of the participant in the Purchase Prompt Condition (N = 75) and 65.3% participants in the Purchase +Appraisal Prompt condition (N = 75) bought the coffee (Omnibus test: $X^2(3) = 17.10$, $p < .01$). A pairwise comparison between the control condition and the appraisal condition replicated our previous

findings, showing that more people bought the coffee in the presence of the appraisal prompt ($X^2(1) = 10.53$, $p < .01$). More importantly, supporting H4, having a purchase opportunity top of mind (purchase prompt condition) also increased purchase rates over the control condition (pairwise comparison: $X^2(1) = 6.37$, $p = .01$). Lastly, adding a second prompt (purchase + appraisal prompt condition) did not have an additional effect on choice compared to the appraisal condition (pairwise comparison: $X^2(1) = .01$, $p = .91$) or the purchase prompt condition (pairwise comparison: $X^2(1) = .70$, $p = .40$). In other words, either an appraisal prompt or a purchase prompt could similarly improve participants' attitudes about instant coffee and increase purchases, but having both produced no additional effect.

Study 5 showed that reminding people that they will have the opportunity to purchase the product induced the same belief updating as an appraisal prompt. This extends our previous findings in two ways: 1) we identified a naturally occurring situation where consumer may spontaneously update their beliefs following a disconfirming experience without an appraisal prompt; 2) we demonstrated that belief updating following an appraisal prompt caused a commensurate shift in a consequential choice. Note that all of the participants were offered the opportunity to buy the coffee at the end of the experiment, so it is worth considering what is different about mentioning the purchase opportunity during the experience versus a few minutes after. One thing that goes away quickly is affect. The affect that consumers felt while tasting the coffee is no longer present and available as an input to choice a few minutes later if it is not converted to propositional thinking during the experience. This suggests that when *actively* considering a pending purchase, people are likely to pay attention to their own affective reactions during an experience (i.e., asking themselves how much do I like this coffee?) and generate propositional thoughts that influence their existing beliefs.

CONCLUSIONS AND GENERAL DISCUSSION

In a series of studies, we show that disconfirming experiences only change beliefs when consumers are prompted to evaluate such experiences (Studies 1-4). Even a prompt to merely consider

without explicitly reporting or recording one's affective reactions is sufficient to increase the impact of experiences on beliefs (Study 3). These effects can last long after the experience is over (Studies 1 and 2). If there is no appraisal prompt during the experience, memories of that particular experience are largely driven by preexisting beliefs about that type of experience and the surprising quality of the experience is lost (Study 2). Furthermore, we show that strongly held beliefs such as those about a broader category can be too entrenched to be updated based on a single surprising experience even in the presence of an appraisal prompt (Study 4). Finally, we find that reminding people of an imminent purchase opportunity right before a surprising experience functions like an appraisal prompt and encourages belief updating, shifting choice share in the direction of the experience (Study 5).

This paper is among the first that asks under what conditions the bottom-up experience is likely to overcome prior expectations and to influence future beliefs. A long line of research has established that pre-existing beliefs often dominate and impede learning from experiences (e.g., Bowen *et al.*, 1992; Goldstein *et al.*, 2008; Nevid, 1981; Plassman *et al.*, 2008; Seymour and McClure, 2008; Wardle and Solomons, 1994), whereas other research shows that people do sometimes update their beliefs following new experiences (e.g., Bolton and Drew, 1991; Oliver, 1977; Swan, 1977). The current paper identifies a key factor that reconciles these disparate findings: whether people articulate their emotional responses to experiences. Drawing upon the APE model (Gawronski and Bodenhausen, 2006), we show that without a nudge to evaluate the affective reactions to an experience, the ephemeral emotion felt during consumption might not be captured and the surprising nature of the experience may not appear in memories of that particular experience or beliefs related to that experience. However, when nudged to consider their reaction during consumption, people make the effort to engage propositional thinking, and surprising experiences change beliefs.

Real-time appraisal prompts are present in common marketing practices. For example, customer service calls often conclude with a satisfaction survey; rideshare customers are prompted to evaluate their experiences either during or immediately after their trips. However, existing literature in consumer research has largely overlooked how these appraisal prompts (or the lack thereof) may influence

consumer learning. This paper bridges the gap by demonstrating the role of hedonic appraisal prompts at the time of consumption in moderating the impact of top-down and bottom-up influences.

Managerial Implications

Practitioners can extract several lessons from our results, as they not only offer guidance as to why brand or product perceptions may be resistant to change, but also offer a course of action for encouraging belief updating. Our results demonstrate that giving consumers better experiences without an external nudge to appraise the experience is not sufficient to influence their beliefs and their future willingness to purchase. Prompts to appraise need not be heavy-handed or intrusive. We show that a simple nudge to consider privately one's reaction to an experience or a reminder of an imminent purchase opportunity substantially enhances learning. The results of Study 5 suggest that managers could encourage learning from an improved consumer experience by simply highlighting a future purchase opportunity. There are also existing marketing tools to encourage consumer appraisals without having to disrupt an experience. For example, the money-back guarantee predicated on the consumer's own satisfaction might be enough to prompt consumers to think about whether they are indeed satisfied and encourage belief updating. Of course, more direct prompts, like the ubiquitous surveys marketers use to get feedback might also serve to help change consumers' beliefs. Our research suggests that one key element in belief updating is the timing of such instruments. Having consumers respond while the experience is still fresh in their minds, or even ongoing may be critical to enshrining an experience in consumers' beliefs.

Furthermore, we have evidence that consumers do not intuit the effects of appraisal prompts on subsequent judgments and choices. As a result, consumers are unlikely to interpret prompts to evaluate an experience as a persuasive or selling tactic, which suggests that such prompts may not be met with consumers' defensive reactions. Appraisal prompts may therefore be more effective in influencing perception and subsequent behaviors than recognizable persuasion attempts.

Finally, our results can also be interpreted as saying that consumer beliefs, once formed, may be very difficult to change by merely improving experiences. Firms should consider changing brand names associated with a negative belief if they are hoping to persuade consumers based largely on an improved experience. Spontaneous hedonic evaluation of those improved experiences may be the exception rather than the rule, especially if consumers already have a well-formed belief coming into those experiences. Further research could help firms predict when they should abandon a brand name versus keep the name when improving products or consumption experiences.

Future Research

Empirically we find that prompts to appraise an experience during or immediately after have similar effects, presumably because feelings about the experience are accessible in both of these instances. In principle, we believe that if the prompt causes people to think about the experience while their affective reaction to that experience is still accessible, then there will be belief updating. An interesting question for future research to explore is exactly how long after an experience an appraisal prompt would enhance belief updating and how this varies across contexts.

The present research has examined experiences that deviate from expectations (e.g., fat-free cheese that tastes good, instant coffee that tastes good, and a clip from a comedy show that is not very funny) and showed that consumers don't spontaneously update their beliefs following an experience of surprising quality. We argue that our appraisal prompts operate by inducing participants to make the effort to engage propositional thinking which is not otherwise engaged. We believe any factor that makes people pay attention to their affective reactions during an experience and capture these otherwise fleeting positive or negative reactions through propositional thinking would facilitate belief updating and induce learning. Future research could fruitfully explore other factors that may generate belief updating. For example, results from Study 5 showed that reminding people of an imminent purchase is enough to facilitate belief updating presumably because explicit consideration of a pending purchase motivates

people to pay attention to their affective responses toward the sampling experience and to engage in the more effortful propositional thinking to capture their ephemeral affect. Are consumers who try a new experience generally thinking about purchase (e.g., when dining at a new restaurant, visiting an amusement park for the first time, watching a new show, etc.)? And even if the purchase goal is central to the experience, (e.g., test driving a car, tasting a free sample in the supermarket), when do consumers focus on their own affective reactions (as needed for updating in our framework) versus focusing on the details of the experience (akin to the physical traits condition in Study 1) which may distract them from their own feelings and lead to less updating?

Similarly, we think certain experiences might be more likely to drive spontaneous belief updating. For example, extreme experiences may spontaneously draw attention to the experiencer's own feelings and lead to belief updating without any external prompts. If future research found support for this idea, such a result would have implications for how managers should most effectively update consumers' brand beliefs. For example, less numerous but more extreme experiences may have a much greater impact on consumer beliefs than more numerous but less extreme experiences. In general, future research that uncovers more conditions where propositional thinking is engaged naturally to produce spontaneous belief updating would greatly improve the efficacy of using experiences to change consumer perceptions.

Study 3 showed that a single surprising experience could alter beliefs about a particular brand (Jerry Seinfeld's comedy). However, Study 4 showed that a single surprising experience does not always alter beliefs about a brand (Dior), even with an appraisal prompt. This combination of results raises a question about when appraising a single experience leads to belief change. Certain beliefs are more resilient to change, perhaps because of greater familiarity with, knowledge about or more experiences relevant to that belief. In addition, certain experiences could provide stronger versus weaker bases for belief change. For example, in study 3, it could be that the clip was a particularly strong deviation from expectations or that participants did not have much knowledge about or experience with Jerry Seinfeld's comedy. It could also be that participants in study 4 had a lot of knowledge about or experience with Dior products or that the bracelet was not as deviant from expectations as the clip was. Future research can

examine how belief updating depends on the surprisingness of the experience and the strength of the belief.

As mentioned in our introduction, we think consumers are likely to update the belief that is the most accessible at the moment of appraisal. One factor that can influence the accessibility of those beliefs is the wording of the appraisal prompt. Indeed, in our studies 1-3 and 5, we did see a variety of beliefs show updating including those about a specific product, brand, or the product category corresponding to which of these beliefs was mentioned in the appraisal prompt. This suggests specific appraisal prompts could influence the level of belief that is made most accessible, which in turn has an impact on the breadth of belief that gets updated. For example, "how much do you like this comedy show" vs "how much do you like Jerry Seinfeld's comedy" could cause updating of different beliefs. This could be an interesting question for future research with important implications for managers. Any brand would presumably like to improve evaluations of its own brand, but not its competitors (or even under some circumstances, the category), suggesting the need for careful wording of appraisal prompts.

One implication of our results from Study 2 is that while memory for the quality of an experience seems ephemeral in the absence of appraisal prompts, other aspects of the experience may be much better remembered even in the absence of any external prompts. Consistent with this idea, we have some preliminary data suggesting that while the surprising real-time dislike of a bitter chocolate is very poorly recalled, the surprising lack of sweetness is well remembered even without any prompts. It would be interesting to examine whether and when retrospective evaluations can be more accurately reconstructed based on bottom-up memories of relevant aspects of an experience.

We also showed some evidence that individuals have an erroneous belief that learning through a disconfirming experience occurs regardless of prompts to evaluate those experiences. Future research might explore consequences of this belief. For example, consumers might seek out experiences with the erroneous idea that experiences per se are sufficient for forming accurate beliefs even if no special effort is made to evaluate those experiences. Or they may assume that their current beliefs must accurately account for their recent experiences.

Concluding Remarks

This paper provides evidence about when and why surprising experiences sometimes influence consumer beliefs and sometimes are overwritten by those beliefs. It seems many experiences may go unnoticed, leaving no trace in memory and having no effect on attitudes and beliefs because the cognitive effort required to incorporate reactions to those experiences into beliefs is often not exerted. Our evidence further suggests that reactions to an experience can be made to last with a simple nudge to spend a moment considering one's reaction to the experience, giving rise to predictions about when top-down versus bottom-up influences will dominate beliefs and attitudes following a surprising experience.

REFERENCES

Anderson EW, Sullivan MW (1993) The antecedents and consequences of customer satisfaction for firms. *Marketing Science.* 12(2): 125-212.

Ariely D and Zauberman G (2000) On the making of an experience: the effects of breaking and combining experiences on their overall evaluation, *Journal of Behavioral Decision Making*. 13(2): 219-232.

Bechara A, Damasio H, Tranel D, Damasio AR (1997) Deciding advantageously before knowing the advantageous strategy. *Science*. 275(5304): 1293-1295.

Benjamin J (2015) Market research fail: how new coke became the worst flub of all time. Available at: https://www.business2community.com/consumer-marketing/market-research-fail-new-coke-became-worst-flub-time-01256904#we1HfH2OWzb5wpTe.99/ (accessed 10 April 2018).

Biederman I (1972) Perceiving real-world scenes. *Science*. 177: 77-80.

Biehal G. Chakravarti D (1983) Information Accessibility as a Moderator of Consumer Choice. *Journal of Consumer Research*. 10(1): 1–14.

Bolton RN. Drew JH (1991). A longitudinal analysis of the impact of service changes on customer attitudes. *Journal of Marketing.* 55(1): 1-9.

Bowen DJ, Tomoyasu N, Anderson M, Carney M, Kristal A (1992) Effects of expectancies and personalized feedback on fat consumption, taste, and preference. *Journal of Applied Social Psychology*. (22): 1061–1079.

Brown KW, Ryan RM (2003) The benefits of being present: mindfulness and its role in psychological well-being. *Journal of Personality and Social Psychology*. 84(4): 822-848.

Buschman TJ, Miller EK (2007) Top-Down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Scienc*. 315(5820): 1860-1862.

Cacioppo JT, Petty RE, Sidera JA (1982) The effects of a salient self-schema on the evaluation of proattitudinal dditorials: top-down versus bottom-up message processing. *Journal of Experimental Social Psychology.* 18(4): 324-338.

Cadotte ER, Woodruff RB. Jenkins RL (1987) Expectations and norms in models of consumer satisfaction. *Journal of Marketing Research.* 24(3): 305-314.

Cohen CE (1981) Person categories and social perception: testing some boundaries of the processing effects of prior knowledge. *Journal of Personality and Social Psychology*. 40: 441-452.

Darley JM, Gross PH (1983) A hypothesis-confirming bias in labeling effects. *Journal of Personality and Social Psychology*. 44(1): 20-33.

Duckworth KL, Bargh JA, Garcia M, Chaiken S (2002) The automatic evaluation of novel stimuli. Psychological science. Nov;13(6):513-9.

Fazio RH, Olsen MA (2003) Implicit measures in social cognition research: their meaning and use. *Annual Review of Psychology*. 54: 297-32

Fazio RH, Sanbonmatsu DM, Powell MC, Kardes FR (1986) On the automatic activation of attitudes. *Journal of personality and social psychology*. *50*(2): 229.

Florack A, Scarabis M, Bless H (2001) When do associations matter? The use of automatic associations toward ethnic groups in person judgments. *Journal of Experimental Social Psychology*. 37: 518-524.

Gawronski B, Bodenhausen GV (2006) Associative and propositional processes in evaluation: an integrative review of implicit and explicit attitude change. *Psychological Bulletin.* 132(5): 692-731.

Goldstein R, Almenberg J, Dreber A, Emerson JW (2008) Do more expensive wines taste better? Evidence from a large sample of blind tastings. *Journal of Wine Economics.* 3(1): 1-9.

Greenwald AG. Leavitt C (1984) Audience involvement in advertising: four levels. *Journal of Consumer Research*. 11(1): 581–592.

Hoch SJ, Deighton J (1989) Managing what consumers learn from experience. *Journal of Marketing*. 53 (2): 1–20.

Hoch SJ, Ha YW (1986) Consumer learning: advertising and the ambiguity of product experience. *Journal of Consumer Research.* 13(4): 221-233.

Hult GTM, Morgeson FV, Morgan NA, Mithas S (2017) Do Managers Know What Their Customers Think and Why? *Journal of the Academy of Marketing Science*. 45(1): 37-54.

Jones EE, Rock L, Shaver KG, Goethals GR, Ward LM (1968) Pattern of performance and ability attribution: an unexpected primacy effect. *Journal of Personality and Social Psychology*. (10)4: 317-340.

Lee L, Frederick S, Ariely D (2006) Try it, you'll like it: the influence of expectation, consumption, and revelation on preferences for beer. *Psychological Science*. 17(12): 1054-1058.

Mazursky D, Jacoby J (1986) Exploring the development of store images. *Journal of Retailing.,* 62(2): 145-165.

Meyvis T, Ratner RK, Levav J (2010) Why don't we learn to accurately forecast feelings? How misremembering our predictions blinds us to past forecasting errors. *Journal of Experimental Psychology: General.* 139(4): 579-589.

Morewedge CK, Gilbert DT, Wilson TD (2005) The least likely of times: how remembering the past biases forecasts of the future. *Psychological Science*. 16(8): 626 – 630.

Morewedge CK, Huh YE, Vosgerau J (2010) Thought for Food: Imagined Consumption Reduces Actual Consumption. *Science*. 330(6010): 1530-1533.

Nevid JS (1981) Effects of brand labeling on ratings of product quality. *Perceptual and Motor Skills*. 53: 407–410.

Novemsky N, Ratner RK (2003) The time course and impact of consumers' erroneous beliefs about hedonic contrast. *Journal of Consumer Research*. 29(3): 507- 516.

Oliver RL (1977) Effect of expectation and disconfirmation on postexposure product evaluations: an alternative interpretation. *Journal of Applied Psychology.* 62(4): 480–486.

Oliver RL (1980) A cognitive model of the antecedents and consequences of satisfaction decisions. *Journal of Marketing Research.* 17(4): 460-469.

Oliver RL (1999) Whence consumer loyalty?" *Journal of Marketing.* 63: 33-44.

Oliver RL, Burke, RR (1999) Expectation processes in satisfaction formation: a field study. *Journal of Service Research*. 3: 196 – 214.

Oliver RL, DeSarbo WS (1988) Response determinants in satisfaction judgments. *Journal of Consumer Research*. 14(4): 495–507.

Olson JC, Dover PA (1979) Disconfirmation of consumer expectations through product trial. *Journal of Applied Psychology*. 64(2): 179-189.

Palmer SE (1975) The effects of contextual scenes on the identification of objects. *Memory & Cognition.* 3(5): 519–526.

Park, CW, Smith DC (1989) Product-Level choice: a top-down or bottom-up process? *Journal of Consumer Research*. 16(3): 289–299.

Park CW, Iyer ES, Smith DC (1989) The effects of situational factors on in-store grocery shopping behavior: the role of store environment and time available for shopping. *Journal of Consumer Research*. 15(4): 422–433.

Raghunathan R, Irwin JR (2001) Walking the hedonic product treadmill: default contrast and mood-based assimilation in judgments of predicted happiness with a target product. *Journal of Consumer Research*. 28(3): 355–368.

Silver, A (2017) What the failure of new coke can teach us about user research and design. Available at: https://alistapart.com/article/what-the-failure-of-new-coke-can-teach-us-about-user-research-and-design/ (accessed 10 March 2018).

Schkade DA, Kahneman D (1998) Does living in California make people happy? A focusing illusion in judgments of life satisfaction. *Psychological Science*. 9(5): 340-346.

Seymour B, McClure SM (2008) Anchors, scales and the relative coding of value in the brain. *Current Opinion in Neurobiology.* 18(2): 173-178.

Shiv B, Carmon Z, Ariely D (2005) Placebo effects of marketing actions: consumers may get what they pay for. *Journal of Marketing Research*. 42(40): 383-393.

Simonson I, Drolet A (2004) Anchoring effects on consumers' willingness-to-pay and willingness-to-accept. *Journal of Consumer Research.* 31(1): 681–690.

Stangor C, McMillan D. (1992) Memory for expectancy-congruent and expectancy-incongruent information: a review of the social and social developmental literatures. *Psychological Bulletin.* 111(1): 42-61.

Swan JE (1977) Consumer satisfaction with a retail store related to the fulfillment of expectations on an initial shopping trip. *Consumer Satisfaction, Dissatisfaction and Complaining Behavior*. Bloomington: Indiana University School of Business.

Swan JE, Trawick FI (1981) Disconfirmation of expectations and satisfaction with a retail service. *Journal of Retailing*. 57(3): 49-67.

Tse DK, Wilton PC (1988) Models of consumer satisfaction formation: an extension. *Journal of Marketing Research.* 25(2): 204-212.

Varelaa P, Aresb G, Giménezb A, Gámbarob A (2010) Influence of brand information on consumers' expectations and liking of powdered drinks in central location tests. *Food Quality and Preference.* 21(7): 873-880.

Wardle J, Solomons W (1994) Naughty but nice: a laboratory study of health information and food preferences in a community sample. *Health Psychology.* 13(2): 180-183.

Weber R, Crocker J (1983) Cognitive processes in the revision of stereotypic beliefs. *Journal of Personality and Social Psychology*. 45(5): 961-977.

Wilson TD, Lindsey S, Schooler TY (2000) A model of dual attitudes. *Psychological Review*. 107: 101-126.