Buying Beauty for the Long Run: (Mis)predicting Liking of Product Aesthetics

EVA C. BUECHEL

CLAUDIA TOWNSEND

Manuscript in its present form has been invited for a second round of reviews at the *Journal of Consumer Research*. It is currently being revised. Please do not cite or circulate without permission.

Eva C. Buechel (eva.buechel@moore.sc.edu) is Assistant Professor of Marketing, Moore School of Business, University of South Carolina, 1014 Greene Street, Columbia, SC 29208. Claudia Townsend (ctownsend@bus.miami.edu) is Assistant Professor of Marketing, School of Business Administration, University of Miami, 5250 University Drive, P.O. Box 248147, Coral Gables, FL 33124.

Contribution Statement

While previous research has shown that consumers have difficulty predicting future hedonic value, surprisingly little is known about how – and how well – consumers predict liking of sensory stimuli over repeated exposures. Furthermore, the limited existing research on predicted and experienced sensory liking over time seems to have overlooked the important and marketing relevant consumption domain of visual satiation.

The present research fills this gap by examining predicted and experienced liking of product aesthetics over time. We focus on two important product design elements: the intensity of color and the intensity of pattern. We demonstrate that consumers predict greater irritation and faster decrease in liking for more intense colors and patterns and, as a result, shy away from intense product designs when making decisions for long-term use. We also show this intuition to be misguided, leading to errors in predicted utility and sub-optimal decision-making.

By directly manipulating the intensity of stimuli, this research is the first to identify if and how such stimuli characteristics systematically influence predicted and experienced sensory utility over time. Thus, the present research not only contributes substantively by focusing on errors in predictions in a new and important consumption domain (i.e., product aesthetics), but it also offers important theoretical insights into the psychological mechanisms determining predicted and experienced hedonic value over time.

Abstract

We identify a systematic error in consumer preferences for aesthetics. When choosing for long-term (vs. short-term) use, consumers are less likely to select products with high arousal designs (i.e., intense colors or intense patterns). This preference is driven by misguided predictions of faster decrease in liking for high arousal product designs over time because they are expected to become increasingly irritating. Six studies test our conceptualization in the lab and in the field. The first four studies examine expected liking for product designs of varying arousal levels over time and how these expectations influence product design preferences for long-term and short-term use. The last two studies then investigate the accuracy of such consumer expectations by directly comparing predicted versus experienced liking for product designs of varying arousal levels over time. The studies reveal a systematic error in prediction whereby consumers overestimate satiation from high arousal product designs. Managerial and theoretical applications are discussed.

It is becoming increasingly difficult for marketers to differentiate products solely on the basis of function. Consequently, markets are becoming more cognizant of the impact of aesthetic design elements such as color and pattern on decision-making (Hoegg and Alba 2008; Page and Herr 2002; Patrick and Peracchio 2010), even for mundane and utilitarian products (Townsend and Shu 2010;Yamamoto and Lambert 1994). Moreover, the quality of aesthetically-driven decisions depends on how well consumers can predict the intensity of future likes and dislikes for different product designs (Kahneman and Snell 1992). The cars we drive, the toothbrushes in our medicine cabinets, and the couches in our living rooms are all examples of products we purchase one day and then repeatedly see for months or even years to come. Thus, for any durable good, consumers must not only identify their current aesthetic preferences, but also predict if and how these preferences may change over time (Kahneman and Snell 1992).

Prior research suggests that consumers have difficulty predicting hedonic value (Wilson and Gilbert 2003, 2005). In general, consumers overestimate the impact of future events and have erroneous beliefs about how liking changes over time (Kahneman and Snell 1992; Snell, Gibbs, and Varey 1995; Wilson and Gilbert 2003, 2005). Yet, research comparing predicted and experienced liking of sensory stimuli is surprisingly inconsistent and sparse. It has not yet examined how consumers predict (vs. experience) liking of visual stimuli over time. Furthermore, it has not systematically manipulated sensory stimulus characteristics to examine how these might differently influence predicted versus experienced hedonic value over time.

Contributing to previous findings both substantively and theoretically, the present research seeks to understand how consumers predict their liking of product designs over time and explore how these predictions may (mis)guide purchase decisions. We focus on two common and important product design elements: the intensity of color and the intensity of pattern.

Specifically, we examine how their arousal potential (i.e., their intensity) differently influences predicted and experienced liking and how these differences impact the nature and quality of consumer decision-making. We posit that consumers choose simple (i.e., low arousal potential) over intense (i.e., high arousal potential) design elements when making decisions for long-term use because they anticipate that intense designs will lead to greater irritation and, as a result, greater decrease in liking over time. We argue that these predictions are often misguided. Contrary to consumers' intuition, we theorize that experienced liking decreases at a lower rate for intense product design elements than for simple ones, leading to errors in predicted utility and suboptimal decision-making.

In the following literature review, we first outline how repetition influences subjective experiences of stimuli in general, while also providing an overview of prior work comparing predicted and experienced hedonic value over time. We then hone in on the present research topic of visual stimuli and how differences in their arousal potential might differently influence liking of product aesthetics in prospect and in experience.

DYNAMICS OF HEDONIC VALUE

Prior research shows that the subjective experience of a stimulus (i.e., its hedonic value) can change with repeated exposure to the stimulus. Subjective hedonic intensity (i.e., the strength of the psychological response) of a stimulus can decrease (adaptation, habituation, or desensitization) or increase (sensitization; Frederick and Loewenstein 1999) over time. For most stimuli and circumstances, adaptation is the norm. In fact, people seem to adapt to many positive and negative stimuli and circumstances, including smells, pains, wealth, incarceration, and even paraplegia (for a review see Frederick and Loewenstein 1999). Importantly, however, there are also instances where adaptation is absent or where the subjective intensity of a stimulus increases over time (i.e., sensitization; Weinstein 1982). For example, individuals have been shown to have difficulty adapting to traffic noise (Weinstein 1982) and exhibit sensitization to dormitory noises (Weinstein 1978) and sexual stimuli (Heiman 1977; Laan, Everaerd, and Evers 1995).

Important to our investigation is that consumers are not very adept at predicting future hedonic value (Wilson and Gilbert 2003, 2005) and how it changes over time (Kahneman and Snell 1992; Snell, Gibbs, and Varey 1995). A vast stream of research in affective forecasting suggests that people overestimate how much and for how long future events – ranging from a positive HIV-test to tenure decisions – will impact their happiness, presumably because they underestimate adaptation to these events (Gilbert et al. 1998; Sieff, Dawes, and Loewenstein 1999; Wang, Novemsky, and Dhar 2009; Wilson and Gilbert 2003, 2005).

Fewer studies have directly compared predicted and experienced liking for *sensory stimuli* over time – the specific topic of the present research. And while the existing studies on sensory liking also demonstrate struggles in predictions, the types of errors are different and less systematic across studies in this domain. Kahneman and Snell (1992) investigated predicted and experienced liking of isolated gustatory and auditory stimuli. They found that while participants correctly anticipated that they would adapt to the taste of ice cream and the sound of music, they overestimated how much they would adapt to the two stimuli, thus *overestimating* adaptation. For other stimuli, Kahneman and Snell (1992) found participants failed to predict even the direction of how hedonic value would change over time. Participants anticipated that the tastiness of an unpalatable yogurt would grow increasingly unpleasant, while experience ratings revealed the taste to become increasingly pleasant. In other words, participants predicted sensitization, but experienced adaptation. Research in the area of sensory hedonic value has also found that consumers fail to intuit that inserting breaks into auditory and somatic experiences will disrupt adaptation, thus worsening negative and improving positive ones (Nelson and Meyvis 2008), which leads to suboptimal consumption decisions that fail to maximize utility in the long run (Nelson and Meyvis 2008; Simonson 1990). In short, these limited past findings make it clear that consumers have difficulty predicting hedonic value for different types of sensory stimuli (e.g., auditory and gustatory) and, as a result, struggle when making decisions about future consumption. However, the findings reveal little consistency in the type or direction of forecasting errors, and at least two important questions remain.

First, despite its marketing relevance, little is known about how – and how well – consumers predict responses to *visual* stimuli over time, and how this might influence consumer decision-making. And because previous research on predicted and experienced liking for sensory stimuli in other domains (e.g., auditory and gustatory) reveal little consistency in the direction or type of error made, extrapolation to visual stimuli is difficult.

Second, research comparing predicted and actual liking in the domain of sensory experiences has yet to systematically manipulate stimuli characteristics and examine how tractable variations may influence predicted and experienced hedonic value over time. Specifically, while mispredictions have been established for isolated auditory (e.g., music and noises) and gustatory stimuli (e.g., ice cream and yoghurt), researchers did not systematically manipulate characteristics of these stimuli (e.g., loudness of sounds or intensity of flavor). It is therefore unclear whether and how predictions about hedonic value vary across stimuli characteristics and how they might influence forecasting errors.

To fill these substantive and theoretical gaps, we manipulate arousal potential (i.e., the intensity) of visual design elements and examine how they might differently influence predicted and experienced liking over time.

DYNAMICS OF VISUAL HEDONIC VALUE FROM PRODUCT DESIGN ELEMENTS

In the present research, we focus on predicted and experienced liking of two common product design elements: color and pattern. These two design elements are important for two reasons. First, they have a strong impact on overall beauty and aesthetics, an affect-based criterion by which consumers evaluate and differentiate between products when making purchase decisions (Alba and Williams 2012; Hoegg and Alba 2008; Page and Herr 2002; Patrick and Peracchio 2010; Postrel 2003). Second, and specifically relevant to this investigation, both color and pattern are aesthetic characteristics that influence complexity and arousal, such that intense designs (i.e., brighter and more saturated colors; more complex patterns) lead to greater perceived complexity and higher psychological and physiological arousal[1] (Baker and Franken 1967; Berlyne 1970; Berlyne and McDonnel 1965; Holtzschue 2011; Küller, Mikellides, and Janssens 2009; Nicki and Gale 1977; Zielińki 2016). For example, when in a colored and patterned room (vs. gray room; Küller et al. 2009) or when shown more (vs. less) colorful or patterned art (Nicki and Gale 1977), participants not only rate the stimuli as more complex, but also show lower alpha rhythm and greater desynchronization in EEG recordings, both indications of neurological arousal. Importantly, these patterns extend to self-reported psychological measures. Research consistently finds that brighter or more saturated colors as well as high contrast and complex stimuli are rated higher on both explicit and indirect measures of arousal (i.e., they are rated as more arousing, more anxiety-producing, more upsetting, and

[1] Two important dimensions of color are the hue (wavelength) and the chroma (saturation, brightness), whereby lower chroma is associated with less pure colors (i.e., pastels or dark colors). According to our conceptualization, arousal potential is mostly determined by chroma, whereby arousal potential is higher for highly saturated colors. But it is important to highlight that within hue – thus keeping saturations constant – there can be differences in arousal potential such that longer wavelengths (e.g. red) are associated with higher arousal compared to shorter wavelengths (e.g. blue; Walters, Apter, and Svebak 1982). These differences, however, are much subtler than the more extreme differences in saturation (Küller, Mikellides, and Janssens 2008; Zielińki 2016). Our investigation thus focuses on chroma.

less calming; Janiszewski and Meyvis 2001; Weller and Livingston 1988; Wright and Rainwater 1962; Zielińki 2016).

The fact that color and pattern influence actual and perceived arousal is important to our investigation because we suspect that these stimuli characteristics differently influence predicted versus experienced liking over time, thus leading to forecasting errors. We first discuss how stimuli characteristics and their arousal potential might influence predicted liking, and we then discuss how they might differently influence experienced liking.

Arousal and Predicted Liking of Product Design Elements

While we are not aware of any pertinent research examining dynamics of predicted hedonic value for visual stimuli over time, there is some, albeit limited, research on the predicted hedonic value for other types of sensory stimuli. In general, a survey of past findings examining consumer intuitions about hedonic value over time reveals that expectations of adaptation seem to prevail for most stimuli (Frederick and Loewenstein 1999; Snell et al. 1995). Importantly, though, a closer look reveals that decision-makers seem to anticipate sensitization to complex and highly arousing stimuli. When examining consumer intuitions for different types of noise, for example, Snell and colleagues (1995) found that participants expect adaptation for low arousal noises (e.g., noise from a highway), but expect sensitization for high arousal noise (e.g., a loud stereo, high-pitched noise, or a pneumatic drill). Accordingly, Nelson and Meyvis (2008) found that consumers expected an annoying noise to become increasingly irritating, such that they would feel more irritated after listening to it for 40 seconds versus 5 seconds.

Based on this suggestive evidence from other sensory domains, we propose that consumers anticipate high arousal design elements (intense colors and patterns) to become increasingly arousing (i.e., sensitization), such that they become aversive and irritating (see

Nelson and Meyvis 2008). We expect that this anticipated irritation, in turn, leads consumers to anticipate faster decrease in liking (i.e., faster satiation[2]) from high arousal design elements than from low arousal potential design elements. More formally:

H1a: Consumers predict faster decrease in liking (i.e., faster satiation) from high arousal design elements (intense colors and patterns) compared to low arousal design elements.

H1b: Consumers predict increased irritation (i.e., sensitization) with high arousal design elements over time as compared to low arousal design elements.

H1c: Predicted irritation will mediate consumer predictions of satiation for high (vs. low) arousal design elements.

Moreover, we expect that these predictions will influence pertinent consumer behaviors and decisions. Specifically, consumers should be more likely to avoid products with high arousal designs when purchasing products for intended long-term (vs. short-term) use[3]. We therefore hypothesize the following:

H2a: Usage length (short-term versus long-term) will influence preferences for products, such that preference for high arousal design elements (intense colors or patterns) versus low arousal design elements will decrease as predicted usage length increases.

H2b: Predicted irritation for high arousal design elements will mediate this effect.

[2] Some research, including research reviewed in this paper, treats satiation as adaptation (i.e., a diminishing positive response to a positive stimulus). We use the term satiation (vs. adaptation) to describe a decrease in liking independent of valence. In other words, in our conceptualization, satiation can refer to both a liked stimulus being liked less over time (i.e., adaptation) and a disliked stimulus being disliked more over time (i.e., sensitization).

[3] Note that our conceptualization is about the amount of exposure (e.g., number of exposures and duration of exposure) over time, and presumably exposure would be greater when products are used over a longer period of time.

This preference, we argue, may be misguided. We next consider literature on experienced liking of visual stimuli, thus informing the validity of consumers' prediction and the resulting decision-making.

Arousal and Experienced Liking of Product Designs Elements

In the domain of perceptual consumption, as in other sensory consumption domains, the intensity of liking of a stimulus generally decreases over time, leading to satiation with the stimulus (Coombs and Avrunin 1977; Redden 2008). Redden (2008), for example, measured enjoyment of a variety of photographs over multiple exposures and found that the enjoyment of the photographs decreased with each exposure. Similarly, children's liking of geometric shapes declines with repeated exposure (Cantor 1968). While this pattern of satiation is prevalent and considered the norm (Nelson and Meyvis 2008), there are instances where the opposite pattern occurs. Extensive research on the mere exposure effects provides evidence of cases where liking of visual stimuli can increase with repeated exposure (Zajonc 1968).

Explaining both of these seemingly contradictory patterns, the direction and the rate of satiation appear to depend on exactly the type of stimuli characteristics that are the focus of this investigation (Cox and Cox 2002). Specifically, Berlyne and colleagues proposed that satiation effects depend on the "arousal potential" of stimuli (Berlyne 1968, 1970; Berlyne and Lawrence 1964; Berlyne and McDonnel 1965). Consumers experience greatest hedonic value when arousal is at a moderate level (Wundt 1974), suggesting that optimal arousal occurs for stimuli with medium arousal potential (Cox and Cox 2002). That is, stimuli should not be too simple, making them boring and tedious, yet also not too complex, raising arousal to an uncomfortable level (Bornstein 1989; Berlyne 1970). Importantly, repeated exposure of stimuli generally decreases their complexity (Cox and Cox 2002). For simple stimuli, repeated exposure quickly moves the stimuli away from the optimal arousal potential, leading to boredom, tedium, and decreased liking (Berlyne 1970). For complex stimuli, repeated exposure moves the stimuli toward optimal arousal potential, thus increasing liking. In line with this, mere exposure effects have been shown to be strongest for unfamiliar and complex stimuli (Bornstein 1989). Only after a large number of exposures does arousal potential decline enough to move past the optimal level of arousal, thus decreasing liking (Zajonc et al. 1972).

While the stimuli and procedures used by Berlyne, Zajonc, and others are relatively contrived, the broad implication for product design is that high arousal design elements might not be as tiring as consumers expect, but instead yield continued pleasure. Moreover, high arousal designs might be less tiring than low arousal designs. Specifically, we hypothesize the following:

H3: Experienced liking decreases at a slower rate (i.e., slower satiation) for high arousal design elements than low arousal design elements.

Together with hypothesis 1a, positing that consumers intuit faster satiation with high arousal design elements than low arousal design elements, we therefore predict that:

H4: Consumers overestimate satiation (i.e., decrease in liking over time) from high arousal design elements, leading to errors in predicted utility.

Overview

We offer support for our hypotheses in six studies. The first two studies reveal that consumers are more likely to avoid products with intense (i.e., high arousal potential) designs when the product is intended for long-term use (hypothesis 2a). This is true when usage duration is directly manipulated (study 1a) and when it is indirectly manipulated using a mindset manipulation (study 1b). We find that this preference for products with low (vs. high) arousal potential designs for long-term use is driven by the greater expected irritation from high arousal potential design elements (hypotheses 1b and 2b). Focusing on hypotheses 1a-1c, the next two studies show that consumers anticipate greater irritation and faster satiation for products with high (vs. low) arousal potential designs (study 2). Accordingly, they expect to use these products less frequently and retain these products for a shorter period of time (study 3). Consistent with our theory and proposed underlying process, we find that these latter two effects are particularly pronounced among individuals who are less comfortable with arousal (study 3). The final two studies (studies 4 and 5) compare predicted and experienced liking of high and low arousal potential design elements to show that consumers overestimate satiation from high arousal potential design elements, leading to errors in predicted utility (hypotheses 3 and 4).

STUDIES 1A AND 1B:
DESIGN PREFERENCES AS A FUNCTION OF CONSUMPTION DURATION

In the first two studies, we examine whether expected consumption duration influences consumer preferences for product design. This allows us to establish downstream consequences of predicted liking over time, namely whether consumers choose less intense product designs (i.e., lower arousal potential; hypothesis 2a) when making decisions for the long run because they anticipate growing irritated with more intense colors and patterns (hypotheses 1b and 2b). Study 1a directly manipulated the usage period, while study 1b used a mindset manipulation to alter the perceived usage period.

Study 1a: Method

Participants consisted of 211 U.S. Mturk workers (46.9% female, $M_{age} = 35.48$; $SD = 12.12$) who were randomly assigned to a short-term or long-term usage condition. Participants were asked to imagine they had decided to purchase paper cups for use over one weekend (short-term condition) or over the coming year (long-term condition).

They were then presented with a choice between two cups, one with a low arousal potential design (white) and one with a high arousal potential design (one of seven variations: solid bright green, solid bright orange, solid bright blue, green stitch, blue dots, black zig-zags, red checkered), see appendix A for stimuli images used in all reported studies. The order of the cup presentation was counterbalanced. Note that only the high versus low arousal manipulations were of theoretical interest. We sampled a variety of high arousal product designs to ensure generalizability.

After making their choice, participants rated their anticipated irritation for the high arousal item (target item), followed by their anticipated irritation for the low arousal item. Specifically, they indicated how the cups would make them feel during the anticipated consumption period (1 = Calm, 7 = Irritated; see Meyvis and Nelson 2008; Zielińki 2016). As a manipulation check, participants then rated the arousal potential of each cup (order counterbalanced), namely whether the cups were intense, novel, colorful, complex, and difficult to process (1= Not at all, 7 = Very much; adapted from Berlyne 1960; white cup: $\alpha(5) = .88$; high arousal cup: $\alpha(5) = .62$). Finally, participants estimated how long they would use the cups (1= Very short period of time, 7 = Very long period of time).

Results

Manipulation Checks. A within-subject ANOVA revealed that the arousal potential index was greater for the colored/patterned cups ($M = 3.79$, $SD = 1.14$) than for the white cup ($M = 1.65$, $SD = 1.08$; $F(1, 205) = 570.09$, $p < .001$, $\eta^2 = .74$). A between-subject ANOVA revealed a significant main effect of consumption duration on estimated usage time ($M_{\text{long-term}}$ = 6.03, *SD* = 1.51; $M_{\text{short-term}}$ = 1.94, *SD* = 1.51; $F(1, 205) = 375.44$, $p < .001$, $\eta^2 = .65$).

Choice. Consistent with hypothesis 2a, participants in the long-term condition were less likely to choose the high arousal cup (43.5%) than participants in the short-term condition (56.5%; $\chi2$ (215) = 4.43, $p = .035$). The results were directionally consistent across design variations when analyzed individually.

Irritation. As predicted (hypothesis 1b), a between-subject MANOVA with usage duration as a factor revealed that for the high arousal cup, participants anticipated being more irritated in the long term condition ($M = 3.78$; $SD = 1.70$) than in the short term condition ($M = 3.29$, $SD = 1.55$); $F(1, 209) = 4.83$, $p = .03$, $\eta^2 = .02$). For the low arousal cup, there was no effect of usage duration ($M_{\text{long-term}} = 2.68$, $SD = 1.58$; $M_{\text{short-term}} = 2.62$, $SD = 1.63$; $F < 1$).

Mediation. To test whether choice was driven by the anticipated irritation with the high arousal cup (hypothesis 2b), we used model 4 of the PROCESS macro (Hayes 2013). Duration was used as the predictor variable, irritation for the high arousal cup (target item) and the low arousal cup as mediators, and choice as the dependent variable. The analysis revealed that the effect of purchase duration on choice was driven by anticipated irritation with the high arousal cup (indirect effect, B = -.22, SE = .13; 95% CI = -.52 to -.02, CI excludes zero; Zhao, Lynch, and Chen 2010), but not the low arousal cup (indirect effect, B = -.03, SE = .10; 95% CI = -.16 to .21). Note that mediation persisted when anticipated irritation with the high arousal cup was entered alone (indirect effect, B = -.18, SE = .10; 95% CI = -.44 to -.03) and when the difference between anticipated irritation with the high and the low arousal cup was used as the mediator (marginal indirect effect, B = -.18, SE = .14; 90% CI = -.49 to -.01).

Study 1b: Methods

Participants consisted of 96 U.S. Mturk workers (63.5% female, M_{age} = 32.85; SD = 9.71) who were randomly assigned to either a short-term or a long-term mindset condition. Mindset was induced using a writing task whereby participants were either asked to write about their plans for the coming week (short-term condition) or the next five years (long-term condition). Specifically, we asked them to list five things that they were planning to do in that time frame. A pretest (N = 54) revealed that this manipulation influenced estimates of the product usage length, such that participants rated usage time (1= Very short period of time, 7 = Very long period of time) to be shorter when they wrote about the near future (M = 2.70, SD = 1.68) than when they wrote about the distant future (M = 4.04, SD = 1.63; $F(1, 51) = 8.56$, $p = .005$, $\eta^2 = .14$).

They were then asked to imagine they had purchased an iPad and were selecting a cover for their new gadget. Participants were presented with two iPad covers, one with a low arousal potential design (K; one of four designs: black, white, grey, or black with blue band) and one with a high arousal potential design (W; one of four designs: bright green, abstract circle, abstract colored, or color palette), see appendix A. The iPad cover combinations and presentation order were randomized across participants. Participants then indicated which of the iPad covers they were more likely to buy on a bipolar scale (1 = Definitely more likely to purchase cover K, 11 = Definitely more likely to purchase cover W). Using the same scale as in study 1a, participants then rated their anticipated irritation during the consumption period for the high arousal item (target item), followed by their anticipated irritation for the low arousal item.

Stimuli. A pretest (N = 250 U.S. Mturk workers) was conducted to select stimuli. Participants rated one of 14 iPad covers on arousal potential (intense, novel, colorful, complex, and difficult to process; 1 = Not at all, 7 = Very much; $\alpha(5) = .80$) and liking (1= Not at all, 7 =

Very much). Eight covers were selected, such that the high and low arousal options varied significantly on arousal potential index ($Ms_{\text{high}} > 3.42$, $Ms_{\text{low}} < 2.38$; ts > 3.12, ps < .004), but did not vary significantly from each other on liking ($Ms_{\text{mostliked}} = 4.33$, $Ms_{\text{leastliked}} = 3.44$; ts < 1.50, ps >.14).

Results

Purchase Likelihood. Consistent with hypothesis 2a, participants in the long-term use condition reported being less likely to purchase the high arousal covers ($M = 3.74$; $SD = 3.39$) than participants in the short-term use condition ($M = 5.64$, $SD = 3.92$; $F(1, 94) = 6.39$, $p = .01$, $\eta^2 = .06$). The results were directionally consistent across design variations when analyzed individually.

Irritation. As predicted in hypothesis 1b, a between-subject MANOVA revealed that participants anticipated greater irritation with the high arousal covers in the long term condition ($M = 4.54$; $SD = 1.66$) than the short term condition ($M = 3.74$, $SD = 1.58$; $F(1, 94) = 5.84$, $p = .02$, $\eta^2 = .06$). For the low arousal items there was no effect of usage duration ($M_{\text{long-term}} = 2.11$, $SD = 1.30$; $M_{\text{short-term}} = 2.32$, $SD = 1.47$; $F < 1$).

Mediation. Supporting hypothesis 2b, a mediation model analogous to the one used in study 1a revealed that the effect of mindset on preference was driven by anticipated irritation with the high arousal iPad cover (indirect effect, B = -.93, SE = .42; 95% CI = -1.92 to -.27), but not with the low arousal cover (indirect effect, B = -.003, SE = .10; 95% CI = -.21 to .21). Again, mediation persisted for anticipated irritation with the high arousal cover when it was entered alone (indirect effect, B = -.93, SE = .39; 95% CI = -1.73 to -.19) and when the difference between anticipated irritation with the high and low arousal cover was used as the mediator (indirect effect, B = -.65, SE = .35; 95% CI = -1.52 to -.07).

Discussion

When considering purchase for long-term (vs. short-term) use, consumers were less inclined to select products with high (vs. low) arousal potential designs, supporting hypothesis 2b. This pattern was observed when usage time was explicitly and directly manipulated (study 1a), as well as when usage time was indirectly manipulated using a mindset manipulation (study 1b). The effect generalized across a variety of low and high arousal designs and across two different product categories (i.e., cups and iPad covers, the latter being a personal and isolated product, thus diminishing concerns about how other users might evaluate the design or how it might match other products). Mediation analysis revealed that the preference for low arousal designs when deciding for the long-term use was driven by greater anticipated irritation with high arousal product designs, supporting hypotheses 1b and 2b.

STUDIES 2 AND 3: PREDICTED LIKING AND PREDICTED USAGE OVER TIME

Having established the downstream consequences of predicted utility over time in studies 1a and 1b, the next two studies more directly and more closely examine consumer expectations of faster decrease in liking for high (vs. low) arousal potential designs (hypotheses 1a-1c). They do this by examining anticipated product liking, anticipated irritation, and anticipated product usage for high (vs. low) arousal designs over time.

Study 2: Predicted Liking and Irritation over Time

In study 2, we directly examine hypotheses 1a -1c by manipulating the arousal potential of product designs and examining the effect on predicted liking and predicted irritation over time. Predictions were made for two product categories generally viewed on a daily basis: plates and bedding sets, with varying product designs. We expected that, across both product categories, participants would predict faster decrease in liking (i.e., satiation) and faster increase in irritation (i.e., sensitization) for high arousal potentials designs than for low arousal designs. We also expected predicted irritation to mediate predicted satiation for high arousal designs, but not for low arousal potential designs.

The manipulation of the arousal potential of the product designs differed from previous studies in two key ways. First, because they are theoretically equivalent (and tend to vary in tandem in the real-world), the first two studies included manipulations of high arousal potential due to both color and pattern, sometimes within the same stimulus. In this study, we now manipulate arousal potential from color and pattern orthogonally to ensure that our proposed effects hold across both design elements. Second, while studies 1a and 1b used actual preexisting products as stimuli to ensure external validity, for this study we created our own stimuli. This enabled us to more systematically manipulate the product designs, thus offering greater control over the key constructs (i.e., arousal potential of designs) in an effort to maximize internal validity.

As a final difference, in contrast to studies 1a and 1b, participants in this study only evaluated one product design at a time (rather than viewing and indicating a preference between two product designs). This ensures that the effects are not due to contrast effects or joint evaluation modes.

Methods

Participants consisted of 300 U.S. Mturk workers (46.9% female, M_{age} = 35.48; *SD* = 12.12) who were randomly assigned to two of 10 conditions in a 5 (product design: low arousal potential [plain white], high arousal potential pattern [striped white or abstract white], high arousal potential color [strong green or strong orange]) x 2 (within factor product category: bedding and plates) nested mixed-subjects design. Note that, again, only the high versus low arousal manipulations were of theoretical interest. Product and design variations were included to ensure generalizability.

Participants were instructed to imagine they had recently moved into a furnished apartment for the foreseeable future. They were then shown a product that they would be using for the time they lived in the apartment. In the bedding conditions, participants were shown an image of a bed with bedding. In the plate conditions, participants were shown an image of a plate that was said to be part of a larger set. The products were presented in one of the five condition-consistent product designs (plain white, striped white, abstract white, strong green, or strong orange), see appendix A. Note that participants saw one bed and one plate (in randomized order), and that the design was randomly determined for each product (i.e., it was unlikely that participants saw the same design for the two different products). Because of this, the responses for the bed and plate were stacked, leading to 600 observations.

For each product, participants were asked to indicate how much they liked the displayed product at the very moment (t1; 0 = Dislike extremely, 100 = Like extremely). Next, participants were asked to imagine using the product for the foreseeable future and to indicate their liking relative to t1 at six points in the future (in one week (t2), one month (t3), six months (t4), one year (t5), two years (t6), four years (t7; -50 = much less, +50 much more).

The procedure was then repeated for the potential mediator measure of irritation. Participants were asked to indicate to what extent the displayed product made them feel irritated at that very moment (t1; 0 = Not at all irritated (Calm), 100 = Very irritated). Next, participants were asked to imagine using the product for the foreseeable future and to indicate their expected irritation relative to t1 at the same six points in the future (-50 = Much less, +50 = Much more).

Stimuli. A pretest (N = 278 U.S. Mturk workers) was conducted to test the perceptions of product designs. Participants were randomly presented with one of the five product designs and rated it on arousal potential (intense, novel, colorful, complex, and difficult to process; 1= Not at all, 7 = Very much; α(5) = .66) and liking (1 = Not at all, 7 = Very much). Simple effects comparing each high arousal design to the low arousal design (i.e., white) revealed that arousal potential was perceived to be greater for each of the high arousal designs ($M_{\text{highest}} = 4.10$, $M_{\text{lowest}} = 3.14$) than for the low arousal design ($M = 1.64$; $ts > 8.22$, $ps < .001$), but liking did not differ for the high arousal designs ($M_{\text{highest}} = 3.23$, $M_{\text{lowest}} = 3.49$) versus the low arousal design ($M = 2.94$; $ts < 1.04$, $ps > .30$).

Results

Predicted liking

Participants' current liking (t1) represented absolute liking at T1. Absolute liking ratings for T2-T7 were calculated by adding the indicated relative change in liking to T1 (T1+t2 = T2, T1+t3 = T3, etc.).

To examine the effect of arousal potential from color and from pattern independently, we compare the plain white design (low arousal) to the colored designs (high arousal) and to the patterned designs (high arousal) independently. Note that, because they are theoretically equivalent, we report the results collapsed across the high arousal pattern (striped white or abstract white) and the high arousal color (strong green or strong orange) conditions.

A 2 (product: bedding, plates) x 2 (color: low arousal, high arousal) x 7 (time: T1-T7) mixed ANOVA did not reveal a significant product x color x time interaction ($F = 2.01$, $p = .12$). A similar mixed ANOVA with 2 (product: bedding, plates) x 2 (pattern: low arousal, high arousal) x 7 (time: T1-T7) did not reveal a significant product x pattern x time interaction (*F* = 1.35, $p = .23$). We therefore report the results collapsed across products in all analyses.

Color vs. Plain White. A 2 (color: low arousal [white], high arousal [green, orange]) x 7 (time: T1-T7) mixed ANOVA revealed a significant color x time interaction ($F(6, 2340) = 3.20$, $p = .004$, $\eta^2 = .01$), indicating that predicted satiation over time differed as a function of arousal potential. While simple effects [linear contrasts] revealed that both high arousal ($F(6, 1206) = 60.70$, $p < .001$, $\eta^2 = .23$ [$F(1, 201) = 93.77$, $p < .001$, $\eta^2 = .32$]) and low arousal ($F(6, 1134) = 20.13$, $p < .001$, $\eta^2 = .10$ [$F(1, 189) = 30.94$, $p < .001$, $\eta^2 = .14$]) designs were marked by a decline in liking, the color x time interaction reveals that participants predicted a greater decrease in liking (faster satiation) from high arousal designs than from the low arousal design over time, supporting hypothesis 1a; see figure 1a. Note that the results held when green and orange conditions were not collapsed (color x time interaction: $F(12, 2316) = 2.07$, $p = .01$, $\eta^2 = .01$).

Pattern vs. Plain White. A 2 (pattern: low arousal [no pattern], high arousal [stripes, abstract]) x 7 (time: T1-T7) mixed ANOVA revealed a significant pattern x time interaction ($F(6, 2358) = 8.54$, $p < .001$, $\eta^2 = .02$), indicating that predicted satiation over time differed as a function of arousal potential. While simple effects [linear contrasts] revealed that both high arousal ($F(6, 1224) = 70.70$, $p < .001$, $\eta^2 = .26$ [$F(1, 204) = 118.11$, $p < .001$, $\eta^2 = .37$]) and low arousal ($F(6, 1134) = 20.13$, $p < .001$, $\eta^2 = .09$ [$F(1, 189) = 30.94$, $p < .001$, $\eta^2 = .14$]) designs were marked by a decline in liking, the pattern x time interaction revealed that participants predicted a greater decrease in liking (faster satiation) from high arousal designs than from the low arousal design over time, supporting hypothesis 1a; see figure 1a. Note that the results held when striped and abstract conditions were not collapsed (pattern x time interaction: $F(12, 2334) = 4.83$, $p < .001$, $\eta^2 = .02$).

Predicted Irritation

Participants' current level of irritation (it1) represented absolute irritation at iT1. Absolute irritation for iT2-iT7 was calculated by adding the indicated relative change in irritation to iT1 (iT1+it2 = iT2, iT1+it3 = iT3, etc.).

Again, we compare the plain white design to the colored designs and the patterned designs independently. Mixed ANOVAs did not reveal any product x color x time ($F = 1.38$, $p > .20$) or product x pattern x time ($F < 1$) interactions for the irritation measures. We therefore again report the results collapsed across products.

Color vs. Plain White. A 2 (color: low arousal [white], high arousal [green, orange]) x 7 (time: T1-T7) mixed ANOVA revealed a significant color x time interaction ($F(6, 2340) = 3.77$, $p = .001$, $\eta^2 = .01$), indicating that predicted irritation over time differed as a function of arousal potential. Participants predicted their irritation to grow more rapidly over time for high arousal designs than for the low arousal design, supporting hypothesis 1b; see figure 1b. Accordingly, simple effects [linear contrasts] revealed a significant effect of time for the high arousal designs ($F(6, 1212) = 13.92$, $p < .001$, $\eta^2 = .06$ [$F(1, 202) = 17.73$, $p < .001$, $\eta^2 = .08$]), but no significant effect of time for the low arousal design ($F(6, 1128) = 1.48$, $p > .18$ [$F(1, 188) = 1.03$, $p > .30$]). Note that the results held when green and orange conditions were not collapsed (color x time: $F(12, 2316) = 2.11$, $p = .01$, $\eta^2 = .01$).

Pattern vs. Plain White. A 2 (pattern: low arousal [white], high arousal [stripes, abstract]) x 7 (Time: T1-T7) mixed ANOVA revealed a significant pattern x time interaction ($F(12, 2346) = 6.76$, $p < .001$, $\eta^2 = .02$), indicating that predicted irritation over time differed as a function of arousal potential. Forecasters predicted their irritation to grow more rapidly over time for high arousal designs than for the low arousal design, supporting hypothesis 1b; see figure 1b.

Accordingly, simple effects [linear contrasts] revealed a significant effect of time for high arousal designs ($F(6, 1218) = 19.35$, $p < .001$, $\eta^2 = .09$ [$F(1, 203) = 26.16$, $p < .001$, $\eta^2 = .11$]), but no significant effect of time for the low arousal design ($F(6, 1128) = 1.48$, $p > .18$ [$F(1, 188) = 1.03$, $p > .30$]). Note that the results also held when striped and abstract conditions were not collapsed ($F(12, 2322) = 3.75$, $p < .001$, $\eta^2 = .02$).

FIGURE 1: RESULTS OF STUDY 2

Figure 1a: Predicted Liking

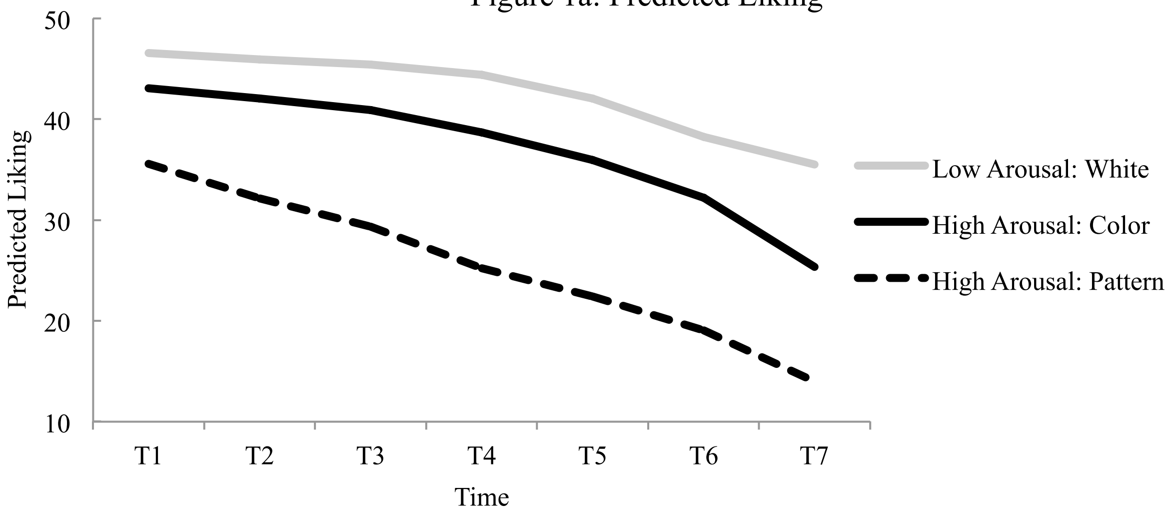
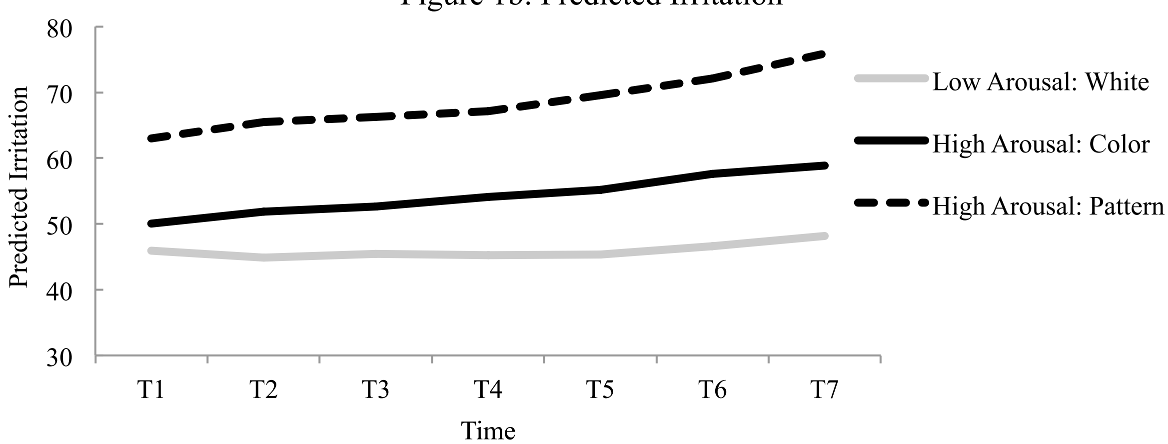


Figure 1b: Predicted Irritation



Moderated Mediation. To test for mediation, we stacked all liking and irritation measures for each respondent and created a variable for time (1-7) to indicate which level of the within subject factor of time it represents. In other words, for each participant and both products, we treated each time period as an observation, leading to 14 observations per participant (each participant provided 7 liking/irritation ratings for a bed and for a plate). In addition, we created a dummy variable to indicate the arousal potential of the design (0 = low arousal potential [white], 1 = high arousal potential [striped, abstract, orange, green]. In line with the pattern of the irritation measures, we expected anticipated irritation to mediate predicted satiation for high arousal potential designs, but not for the low arousal potential design (hypothesis 1c). To test this, we used model 8 of the PROCESS macro (Model 8; Hayes 2013). Time was used as the predictor variable, arousal potential as the moderator variable, anticipated irritation as the mediator variable, and liking as the dependent variable. In addition, product (bed, plate) and respondent ID were entered as covariates.

In line with hypothesis 1c, for the high arousal potential designs, the pathway from time to liking through anticipated irritation was significant and did not include zero (indirect effect, B = -.68, SE = .15; 95% CI = -.98 to -.41), supporting mediation (Zhao et al. 2010). The same was not observed for the low arousal potential design. For the low arousal potential design conditions, the pathway from time to liking through anticipated irritation included zero (indirect effect, B = -.14, SE = .17; 95% CI = -.52 to .19), not supporting mediation. Note that in the interest of brevity, we tested arousal from pattern and arousal from color together in this mediation analysis. However, the results held when conducting separate mediation analyses for the pattern and color conditions. Results also persisted when dropping the covariates (product and respondent ID) from the analyses.

Discussion

Participants expected to satiate more quickly (a faster decrease in liking) from high (vs. low) arousal potential product designs because they anticipated growing increasingly irritated with high arousal product designs over time, thus providing direct evidence for hypotheses 1a-1c. Importantly, we separately and systematically varied color and pattern in this study to ensure that our effect holds independent of whether the arousal stems from color or from pattern.

A caveat of this study is that, despite having been pretested to be liked equally, the initial liking of the patterned designs was lower than the white and colored designs ($ts > |2.8|$, $ps < .04$), potentially raising the concern that consumers simply predict faster satiation for less liked designs. Mitigating this concern, the effect was observed for high arousal from color, even though initial liking for white and colored designs did not differ ($t = 1.47$, $p = .14$). Study 4 will rule out this alternative account further.

Having established that consumers anticipate greater satiation from high arousal design, we next explore a variety of expected product usage measures to examine whether this anticipated satiation influences more behavioral, and thus marketing relevant, dependent variables. We also examine a theoretically relevant individual difference as a moderator.

Study 3: Predicted Product Usage over Time

Study 3 examines whether the arousal potential of a product design influences predicted product usage, a behavioral proxy for anticipated satiation. If participants think that they will satiate more quickly from a product with a high (vs. low) arousal design, they should a) anticipate using it less frequently in the future, b) anticipate using it fewer number of times (i.e., cease use after fewer uses), and c) be less likely to predict using it at a distant (vs. a proximate) point in the future than a product with low arousal design.

Importantly, if our theory is correct that this is due to predicted aversive sensitization to arousal (i.e., irritation) from high arousal potential designs over time, leading to faster predicted satiation, then the effect should be moderated by individuals' comfort with arousal. On average, we predict less anticipated usage for products with high (vs. low) arousal designs, in line with previous studies. However, this should be particularly pronounced for individuals who are less likely to enjoy arousal (e.g., people who score low on the Arousal Seeking Scale; Mehrabian and Russel 1973). These individuals should anticipate arousal to become aversive more quickly than individuals who enjoy arousal and therefore be more likely to anticipate lower usage for products with high arousal designs.

Thus, in this study, we manipulated arousal potential of a product (exercise clothing) and measured anticipated product use as well as individuals' Arousal Seeking Tendency (Mehrabian and Russell 1973). We expected that, on average, participants would anticipate lower usage of high (vs. low) arousal product designs, but that this effect of arousal potential on predicted usage would be strongest for individuals who score low on the Arousal Seeking Tendency scale (Mehrabian and Russell 1973).

Methods

Participants were 225 undergraduate students (74.1% female, M_{age} = 20.70, SD = 1.96) who were randomly assigned to a condition in a 2 (arousal potential: low [black, gray, gray dots, dark green, navy], high [bright green, green pattern, purple orange pattern, red, neon ombre]) nested between-subject design, with Arousal Seeking Tendency as a measured variable.

Participants were asked to imagine that they had just received the workout shirt displayed on their computer screen. In the high arousal potential design conditions, they were shown one of five shirts, superimposed with a high intensity design (bright green, green pattern, purple orange pattern, red, neon ombre). In the low arousal potential design conditions, they were shown one of five shirts, superimposed with a low intensity design (black, gray, gray dots, dark green, navy); see appendix A. Note that participants who indicated being male at the beginning of the survey saw a men's shirt, while participants who indicated being female saw a more fitted women's shirt.

Participants were then asked multiple questions about their anticipated use of the workout shirt. In a first set of questions, participants were asked: "Thinking about this workout shirt in the context of all of the workout shirts you own, how often would you use this shirt in the future?" (1 = Never, 9 = All the time) and "After how many wears do you think you'll stop wearing it or get rid of it?" (0 - 500 wears). In addition, they were asked: "How likely are you to wear the workout shirt to your next workout?" (t1), and "How likely are you to wear the workout shirt to a workout one year from now?" (t2; 1 = Very unlikely, 7 = Very likely). Participants also rated their current liking of the shirt (1 = Dislike very much, 7 = Like very much). Lastly, participants filled out an abbreviated version of the Arousal Seeking Tendency (AST) scale (original scale: Mehrabian and Russel 1973). This scale contained items such as "I like to look at pictures that are puzzling in some way" or "My ideal home would be peaceful and quiet".

Stimuli. A pretest (N = 255 undergraduate students) was conducted to select the stimuli. Participants rated one of 13 workout shirts, superimposed with different designs, on arousal potential (intense, novel, colorful, complex, and difficult to process; 1 = Not at all, 7 = Very much; $\alpha(5) = .81$) and liking (1 = Not at all, 7 = Very much). Ten shirts were selected, such that high and low arousal stimuli varied significantly on arousal potential index ($M_{high} = 3.93$, $M_{low} = 1.90$; $F(1, 251) = 238.82$, $p < .001$, $\eta^2 = .48$), but did not vary significantly on liking ($M_{high} = 4.19$, $M_{low} = 3.91$; $F(1, 251) < 1.80$, $p > .18$).

Results

Multiple regression analyses examined how arousal potential of the product design (high vs. low), the arousal seeking tendency (AST; continuous), and their interaction influenced predicted product usage in the future.

Predicted Frequency of Usage. Main effects of arousal potential (B = -3.98, SE = 1.53, $t(221) = -2.59$, $p = .01$; 95% CI = -7.01 to -.95) and AST (B = -.87, SE = .43, $t(221) = - 2.04$, $p = .04$, 95% CI = -1.71 to -.03) were qualified by a significant arousal potential x AST interaction (B = .59, SE = .27, $t(221) = 2.14$, $p = .03$, 95% CI = .04 to 1.14), see figure 2a. Further examining this interaction, slope analysis (Aiken and West 1991; Fitzsimons 2008) revealed that arousal potential influenced low AST individuals' (-1SD) usage frequency estimates (B = -1.35, SE = .39, $t(221) = -3.39$, $p < .001$, 95% CI = -2.13 to -.56). In contrast, it did not influence usage frequency estimates for high AST individuals (+1SD; t <1). Results persisted when current liking was used as a covariate in the analyses, suggesting that the results are not driven by differences in liking for the product designs. Instead, predicted frequency of usage seems to be driven by low (vs. high) AST individuals anticipating greater sensitization and satiation for high arousal potential designs.

Predicted Wears until Discarded. Main effects of arousal potential (B = -271.64, SE = 106.07, $t(207) = -2.56$, $p = .01$, 95% CI = -480.78 to -62.50) and AST (B = -68.75, SE = 28.85, $t(207) = -2.38$, $p = .02$, 95% CI = -125.63 to -11.86) were qualified by a significant arousal potential x AST interaction (B = 40.02, SE = 19.21, $t(207) = 2.08$, $p = .04$, 95% CI = 2.13 to 77.91), see figure 2b. Further examining this interaction, slope analysis revealed that arousal potential influenced low AST individuals' (-1SD) predicted number of total wears (B = -94.02, SE = 27.04, $t(221) = -3.47$, $p < .001$, 95% CI = -147.34 to -40.71). In contrast, arousal potential did not influence predicted number of wears for high AST individuals (+1SD; t <1). Again, results persisted when current liking was used as a covariate in the analyses, suggesting that the results are not driven by differences in liking for the product designs, but by differential anticipated sensitization and satiation for high arousal potential designs.

Likelihood of Wearing in One Year. For likelihood of wearing in one year (t2), a main effect of arousal potential (B = -3.54, SE = 1.48, $t(221) = -2.38$, $p = .02$, 95% CI = -6.47 to -.62) and a marginal main effect of AST (B = -.79, SE = .41, $t(221) = -1.93$, $p = .054$, 95% CI = - 1.60 to .01) were qualified by an arousal potential x AST interaction (B = .56, SE = .26, $t(221) = 2.11$, $p = .04$, 95% CI = .04 to 1.09), see figure 3c . Slope analysis revealed that while arousal potential influenced low AST individuals' (-1SD) predicted likelihood of wearing the shirt in one year (B =-1.04, SE = .38, $t(221) = -2.71$, $p = .01$, 95% CI = -1.80 to -.28), it did not influence likelihood estimates for high AST individuals (+1SD; t <1).

Likelihood of Wearing to Next Workout. Importantly, for the likelihood of wearing the shirt to the next workout (t1), there was no main effect of arousal potential, no main effect of AST, and no arousal potential x AST interaction ($|ts| < 1.50$, $ps > .13$). The fact that AST only influenced likelihood of wearing the shirt in the distant future (t2) suggests that the different pattern observed for high and low AST individuals resulted from different rates of sensitization and satiation with the product designs over time, not simply from differential propensities to use high versus low product designs more generally.

FIGURE 2: RESULTS OF STUDY 3

Figure 2a: Frequency of Use

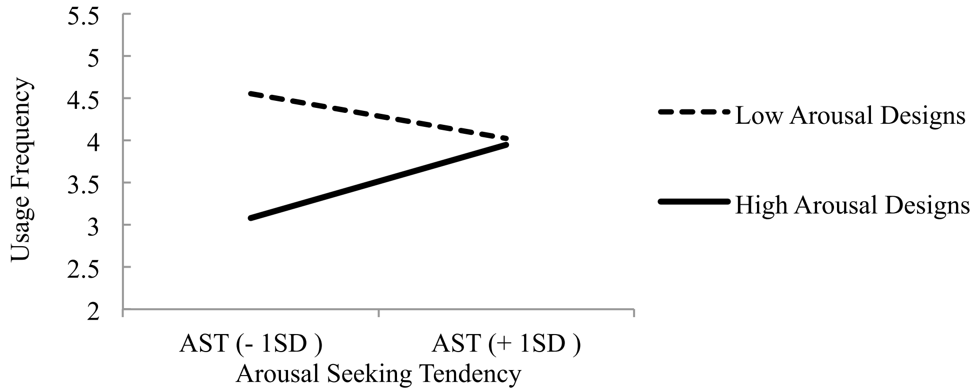


Figure 2b: Predicted Number of Wears (Until Discarded)

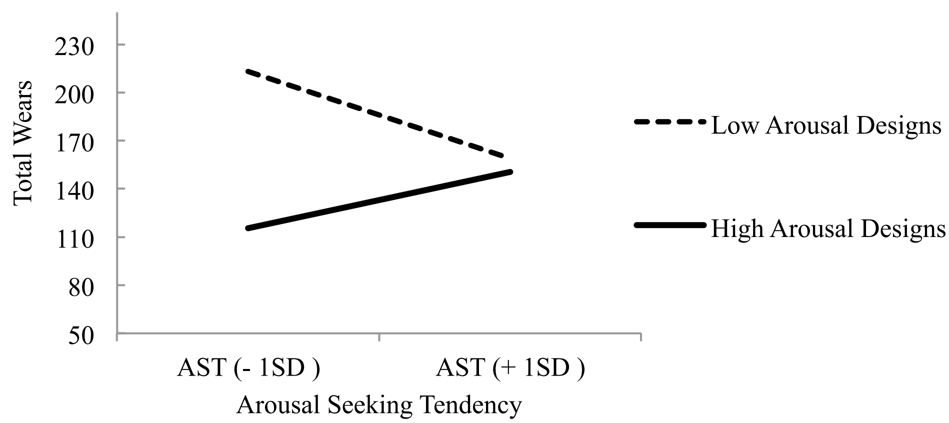
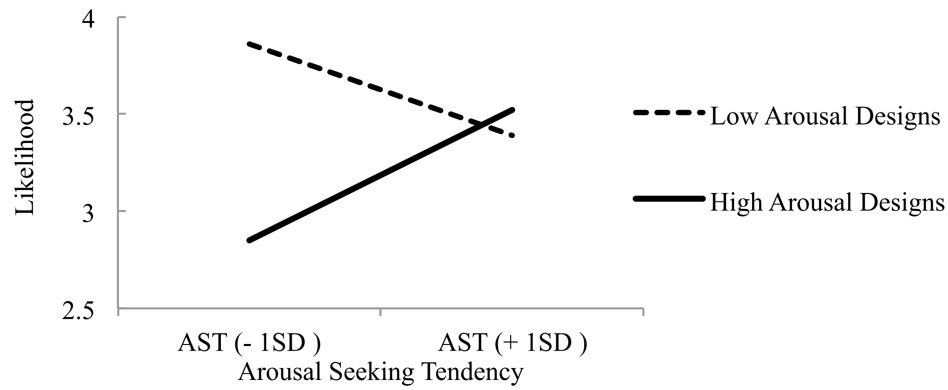


Figure 2c: Predicted Likelihood of Wearing in One Year



Discussion

On average, participants anticipated lower usage of workout shirts (lower frequency and fewer total number of wears) with high (vs. low) arousal potential designs, despite having been pretested and selected to be liked equally. This suggests that, in line with our theory, differential satiation for high and low arousal designs are driving product usage estimates.

More importantly, this study also reveals a theoretically relevant moderator. Arousal potential had greater influence on anticipated product usage for people who scored low on the Arousal Seeking Tendency scale (i.e., who do not enjoy arousal). This pattern persisted even when controlling for liking. Furthermore, arousal potential and AST had greater influence on usage estimates in the distant (vs. near) future. This suggests that arousal potential is indeed the aspect of the stimuli driving predicted product usage over time. Furthermore, and in line with our theory, anticipated product usage seems to be driven by low (vs. high) AST individuals anticipating greater sensitization and faster satiation for high arousal potential designs.

The demonstrated effects are especially compelling given that high arousal items are more unique and less likely to be already owned, thus providing an opportunity for variety seeking. The fact that we observe lower (vs. higher) average usage for high arousal items suggests that our effect is so strong that it overrides the opportunity to seek variety.

Overall then, this study supports our theorizing that, on average, high arousal designs are anticipated to become increasingly arousing and irritating (i.e., sensitization) over time. Using product usage as a behavioral proxy of anticipated satiation, it also shows how these expectations influence pertinent anticipated consumer behaviors. We next focus on whether these intuitions may be misguided.

STUDIES 4 AND 5: COMPARING PREDICTED AND EXPERIENCED LIKING

Thus far, our investigation has focused on predicted satiation for high and low arousal potential product designs and how these predictions influence anticipated product usage and consumer decisions (hypotheses 1 and 2). The remaining two studies examine the soundness of these predictions and decisions.

Contrary to consumers' intuition, we theorized that consumers might satiate less (i.e., liking decreases at a slower rate) from high arousal design elements than low arousal design elements (hypothesis 3). As a result, we hypothesized that forecasters overestimate satiation (i.e., decrease in liking over time) from high arousal design elements (hypothesis 4), leading to errors in predicted utility. To test this, the next studies directly compare predicted and experienced liking of varying product designs over time in both a lab (study 4) and a field setting (study 5).

Study 4: Comparing Predicted and Experienced Liking

Study 4 directly compares predicted and experienced liking of product designs in a laboratory setting. All participants were exposed to a screen background with either a high or a low arousal potential design and reported their liking of the product. Forecasters then predicted their liking over 14 subsequent exposures. In contrast, experiencers were exposed to the product for 14 additional times and reported their liking for each exposure. Replicating previous studies, we expected forecasters to predict greater satiation for high (vs. low) arousal product design. In contrast, we expected experiencers to satiate less from high (vs. low) arousal designs. The discrepancy, we predicted, would lead to forecasting errors whereby forecasters overestimate satiation from high arousal designs.

Methods

Participants were 202 undergraduates and U.S. Mturk workers (46% female, M_{age} = 28.23, *SD* = 12.53) randomly assigned to one condition in a 2 (mode: forecaster, experiencer) x 2 (arousal potential: low [light gray minimal pattern, light green minimal pattern], high [black and white pattern, colored pattern]) nested between-subject design.

All participants were instructed to imagine they were considering a new screen background for their computer. They were then told that, in this study, they would see the screen background 15 times for five seconds to simulate usage over time. Then all participants saw a screen background (in accordance with their assigned arousal potential condition, see appendix A) for five seconds and rated their liking of the screen background on an analog scale (T1; -100 = Dislike extremely, 100 = Like extremely; anchored at 0). After this initial liking rating, forecasters were then asked to imagine the next exposure and predict their liking of the screen background on the same scale (T2), repeating this exercise for each of the 13 future (T3-T15) exposures. After making the predictions, they saw the screen background 14 times for five seconds. Experiencers, by contrast, saw the screen background 14 times for five seconds each. Using the same scale as forecasters, they rated their liking of the screen background after each exposure. To ensure that experiencers viewed each exposure (i.e., did not leave the survey or look elsewhere), they were asked to press a “I am still looking” button displayed below the product image for four of the 14 exposures.

A pretest (N=250 U.S. Mturk workers) was conducted to select stimuli. Participants rated one of 13 screen backgrounds on arousal potential (intense, novel, colorful, complex, and difficult to process; 1= Not at all, 7 = Very much; $\alpha(5) = .76$) and liking (1= Not at all, 7 = Very much). Four screen backgrounds were selected such that they varied significantly on arousal potential index ($Ms_{high} > 3.67$, $Ms_{low} < 1.92$; ts > 5.93, $p < .001$), but did not vary significantly on liking ($Ms_{mostliked} = 3.33$, $Ms_{leastliked} = 2.90$; ts < 1).

Results

Given that the two low arousal potential product designs and the two high arousal potential product designs were theoretically equivalent, the liking ratings were collapsed across these conditions and submitted to a 2 (mode: forecaster, experiencer) x 2 (arousal potential: high, low) x 15 (time: T1-T15) mixed ANOVA. The results revealed a significant mode x time interaction ($F(14, 2772) = 3.95$, $p < .001$, $\eta^2 = .02$), indicating that forecasters, on average, expected to satiate more than experiencers actually did. More important, the results revealed a significant mode x arousal potential x time interaction ($F(14, 2772) = 3.41$, $p < .001$, $\eta^2 = .02$), see figure 3. To further examine the nature of the interaction, we examined how forecasters and experiencers rated the high and low arousal potential designs over time.

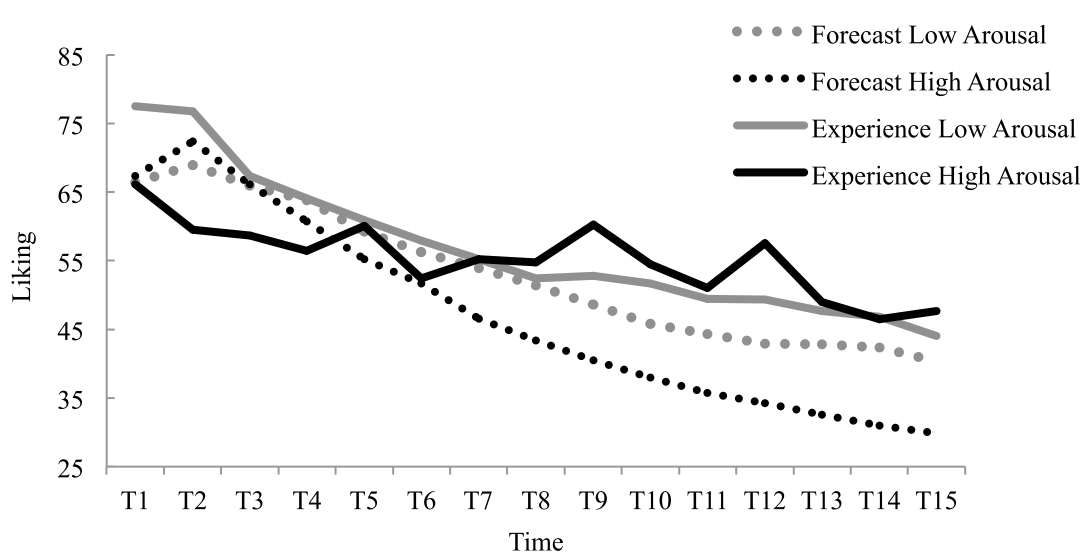
Among forecasters, in line with hypothesis 1a and replicating the findings of study 2, the analysis revealed an arousal potential x time interaction ($F(14, 1246) = 2.22$, $p = .006$, $\eta^2 = .02$), indicating that the predicted satiation over time differed as a function of arousal potential. While simple effects [linear contrasts] revealed that both high ($F(14, 602) = 39.37$, $p < .001$, $\eta^2 = .48$ [$F(1, 43) = 54.76$, $p < .001$, $\eta^2 = .56$]) and low arousal designs ($F(14, 644) = 24.82$, $p < .001$, $\eta^2 = .35$ [$F(1, 46) = 33.33$, $p < .001$, $\eta^2 = .42$]) were marked by a decline in predicted liking, the arousal potential x time indicates that forecasters predicted to satiate more rapidly from high arousal designs than from low arousal design over time.

Interestingly, however, experiencers showed the opposite pattern. Among experiencers, there was also an arousal potential x time interaction ($F(14, 1526) = 2.03$, $p = .01$, $\eta^2 = .02$), indicating that the experienced satiation over time differed as a function of arousal potential.

Contrary to forecasters' predictions, however, simple effects [linear contrasts] revealed that experiencers tired more quickly from low arousal designs (F(14, 672) = 15.18, p < .001, η^2 = .24 [F(1, 48) = 21.69, p < .001, η^2 = .31]) than high arousal designs (F(14, 854) = 1.75, p = .04, η^2 = .03 [F(1, 61) = 3.32, p = .07, η^2 = .05]), thus supporting hypothesis 3.

Examining the accuracy of forecasts, we then examined forecasts and experiences at each level of arousal potential. For screen backgrounds with low arousal designs, there was no mode x time interaction (F < 1), suggesting that forecasters relatively accurately predicted experienced liking over time. For high arousal designs, on the other hand, there was a mode x time interaction (F(14, 1456) = 4.91, p < .001, η^2 = .05), supporting hypothesis 4 and revealing an error in prediction. While simple effects [linear contrasts] revealed an effect of time for both predicted (F(14, 602) = 39.37, p < .001, η^2 = .48 [F(1, 43) = 54.76, p < .001, η^2 = .56]) and experienced liking (F(14, 854) = 1.75, p = .04, η^2 = .03 [F(1, 61) = 3.32, p = .07, η^2 = .05]), the mode x time interaction indicates that forecasters predicted to satiate more rapidly from high arousal potential designs than experiencers actually did.

FIGURE 3: RESULTS OF STUDY 4



Discussion

Replicating the findings of study 2 and supporting hypothesis 1a, forecasters expected to satiate more quickly from high arousal potential product designs than low arousal potential product designs. More important, ratings from experiencers revealed forecasters' intuitions to be misguided. In contrast to forecasters predictions, experiencers satiated more quickly from low arousal designs than from high arousal designs, supporting hypothesis 3. As a result, and supporting hypothesis 4, participants overestimated satiation from high arousal designs.

Note that on average (i.e., averaging across designs), forecasters overestimated satiation from repeated exposure to product designs, which is in line with previously discussed findings from other sensory domains (i.e., music and ice cream; Kahneman and Snell 1992). Importantly, our theory suggests that because anticipated satiation is greater for high (vs. low) arousal stimuli, forecasting errors are more likely or more pronounced for high (vs. low) arousal stimuli. In line with this, forecasters mispredicted liking for high arousal designs over time, but more accurately predicted liking for low arousal designs over time. That being said, our theory does not necessarily preclude forecasting errors to occur for low arousal designs.

Aside from showing errors in prediction, this study also serves to extend and rule out potential alternative explanations in previous studies. Unlike the stimuli used in study 2, liking for low and high arousal potential designs were identical at time 1 ($F < 1$). This further rules out the potential alternative explanation that people simply predict faster satiation from less appealing designs. Furthermore, by including low arousal potential designs involving (albeit minimal) color and pattern, this study generalizes the findings of study 2 to other low arousal designs and further cements the fact that our conceptualization is about the arousal potential of color and pattern and not simply the presence or absence of color and pattern. Lastly, predicted liking was measured on a different scale than in study 2 (in this study participants indicated liking for T1-T15 on the same absolute scale instead of making a relative judgment compared to T1 in study 2), thus demonstrating the robustness of the effect across different elicitation procedures.

This study compared predicted and experienced liking in a relatively contrived experimental setting, whereby exposures occurred in a fairly abbreviated timeframe. In the next study, we use a field experiment to examine whether the discrepancy in predicted and experienced liking holds in a more natural consumption setting.

Study 5: Comparing Predicted and Experienced Liking in the Field

Study 5 extends study 4 by testing hypotheses 3 and 4 in a real consumption domain over a longer period of usage time. In a field study, participants either predicted or reported experienced liking of a product (a name card) over one semester. Again, we expected forecasters to predict faster satiation for products with high (vs. low) arousal potential designs. In contrast, we expected experiencers to satiate faster from products with low (vs. high) arousal potential designs, leading to errors in predicted liking whereby forecasters overestimate satiation from high arousal designs.

Methods

Participants were 306 undergraduates (49.4% Female; M_{age} = 19.90, SD = 1.53) enrolled in ten sections of an introductory business course at a large US university. Participants were assigned to one condition of a 2 (mode: forecaster, experiencer) x 2 (arousal potential: low [plain white or pale green], high [abstract pattern or bright green] nested between-subject design. See appendix A for stimuli images.

Experiencers (N = 122) were students in four sections of a course that met twice a week for 75 minutes over the course of 13 weeks. On the first day of class, participants were randomly given one of four name cards to use during class throughout the semester. The cards either featured a low arousal potential design (plain white or pale green) or a high arousal potential design (abstract pattern or bright green, both already pretested and used in study 2). The name cards with their name on it were distributed at the beginning of every following class and sat on the students' desk for the entirety of each class. Experiencers were given a brief questionnaire at three points in the semester: during the first week (T1), during the sixth week (i.e., before spring break; T2) and during the last week (i.e., before finals, T3). Among other questions related to the course, participants rated how much they currently liked their name card (1= Not at all, 9 = Very much). They also provided a code to match their responses over the three measurement times. Twenty-nine participants were excluded from the analyses because they either had missing data due to absence or because we were unable to match their responses across repeated measures, leaving 93 experiencers.

Forecasters were 184 students enrolled in the six remaining sections of the same class who participated in a survey during the first week of the semester. They were exposed to all four name cards used by experiencers, but they were only asked to evaluate a randomly assigned one. They were asked to imagine having been given the name card on the first day of class to use during class throughout the semester. Using the same scale as experiencers, they were then asked to rate how much they liked their name card (T1). They were then asked to imagine it being shortly before spring break (T2) and predict how much they would like the name cards having used the card in every class. They repeated this exercise for the time before finals (T3).

Results

Given that the two product designs in the high and low arousal conditions were theoretically equivalent, liking ratings were collapsed across arousal potential and submitted to a 2 (mode: forecaster, experiencer) x 2 (arousal potential: high, low) x 3 (time: T1-T3) mixed ANOVA. The results revealed a directional mode x time interaction ($F(2, 546) = 2.19$, $p = .10$, $\eta^2 = .01$), indicating that forecasters, on average, tended to expected to satiate more than experiencers actually did. More important, the results revealed a significant mode x arousal potential x time interaction ($F(2, 546) = 6.71$, $p = .001$, $\eta^2 = .02$), see figure 4. To further examine the nature of the interaction, we first examined how forecasters and experiencers rated the high and low designs over time.

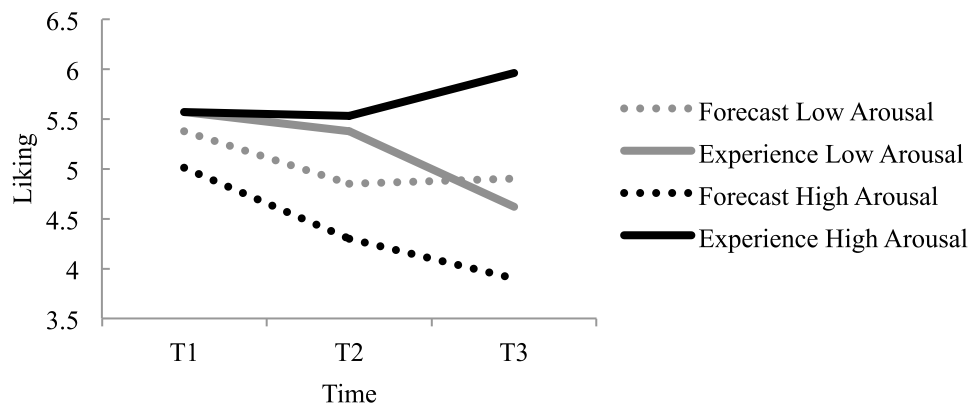
Among forecasters, the analysis revealed an arousal potential x time interaction ($F(2, 364) = 2.85$, $p = .059$), in line with hypothesis 1a and replicating the findings of studies 2 and 4. While simple effects [linear contrasts] revealed that high ($F(2, 202) = 15.65$, $p < .001$, $\eta^2 = .14$ [$F(1, 101) = 21.08$, $p < .001$, $\eta^2 = .17$]) and low arousal potential designs ($F(2, 164) = 4.12$, $p < .001$, $\eta^2 = .02$ [$F(1, 81) = 4.48$, $p < .001$, $\eta^2 = .04$]) were marked by a decline in predicted liking, the arousal potential x time interaction indicated that forecasters predicted more rapid satiation from high arousal designs than from low arousal designs over time.

Experiencers, however, again showed a different pattern. Among experiencers there was an arousal potential x time interaction ($F(2, 182) = 3.88$, $p = .02$, $\eta^2 = .04$), supporting hypothesis 3. Simple effects [linear contrasts] revealed that experiencers tired from low arousal designs over time ($F(2, 82) = 4.64$, $p = .01$, $\eta^2 = .10$ [$F(1, 41) = 6.45$, $p = .02$, $\eta^2 = .14$]), whereas liking ratings for high arousal designs remained constant over time, with no effect of time in this study ($F< 1$ [$F < 1$]).

Examining the accuracy of consumer predictions (hypothesis 4), we then compared forecasts and experiences at each level of arousal potential. For name cards with low arousal designs, there was a marginally significant mode x time interaction ($F(2, 244) = 1.93$, $p = .08$, $\eta^2 = .02$). However, simple effects did not reveal any significant forecasting errors ($F(1, 122) < 1.96$, $p > .16$), suggesting that forecasters relatively accurately predicted experienced liking over time.

For the high arousal designs, there was a significant mode x time interaction ($F(2, 302) = 6.67$, $p = .001$, $\eta^2 = .04$), which revealed the hypothesized error in prediction. Simple effects [linear contrasts] revealed that forecasters predicted they would grow tired of high arousal designs, with a highly significant effect of time on liking ($F(2, 202) = 15.65$, $p < .001$, $\eta^2 = .14$ [$F(1, 101) = 21.08$, $p < .001$, $\eta^2 = .17$]). In fact, liking ratings by experiencers remained more constant over time, with no effect of time ($F < 1$ [$F < 1$]). Furthermore, a between-subject MANOVA examining the effect of mode for each time measurement among high arousal designs revealed that while forecasters and experiencers liked the high arousal name cards equally at the beginning of the semester (T1; $M_{\text{Forecast}} = 5.01$, $SD = 2.49$ vs. $M_{\text{Experience}} = 5.75$; $SD = 1.94$; $F(1, 151) = 1.96$, $p = .16$, $\eta^2 = .01$), forecasters underestimated how much they would like the high arousal name cards at the midpoint of the semester (T2, $M_{\text{Forecast}} = 4.29$, $SD = 2.15$ vs. $M_{\text{Experience}} = 5.53$; $SD = 1.94$, $F(1, 151) = 11.90$, $p = .001$, $\eta^2 = .07$) and at the end of the semester (T3, $M_{\text{Forecast}} = 3.87$, $SD = 2.23$ vs. $M_{\text{Experience}} = 5.69$; $SD = 1.83$; $F(1, 151) = 25.12$, $p < .001$, $\eta^2 = .14$).

FIGURE 4: RESULTS OF STUDY 5



Discussion

Replicating the findings of studies 2 and 4 and supporting hypothesis 1a in a field experiment, forecasters expected to satiate more quickly from high arousal potential product designs over the course of a semester than from low arousal potential product designs. Importantly, experiencer reports reveal this intuition to be misguided. In contrast to forecasters' predictions, experiencers satiated more quickly from low arousal designs than high arousal designs. As a result, they overestimated satiation from high arousal designs, thus underestimating liking of the high arousal design at the mid-point and the end of the semester.

GENERAL DISCUSSION

Product aesthetics are an important determinant of consumer preference and choice for a wide range of products (Hoegg and Alba 2008; Page and Herr 2002; Patrick and Peracchio 2010; Townsend and Shu 2010; Yamamoto and Lambert 1994). Importantly, for any durable product that is purchased for use over an extended period of time, consumers must not only judge their current liking of the product, but also predict if and how their liking of the product may change over time. This research examined how people predict and experience liking for two important design elements – color and pattern – and how such predictions might (mis)guide consumer decision making.

Six studies examined how the arousal potential of product design elements (intensity of color and pattern) differently affects predicted and experienced liking over time and how this influences the nature and quality of consumer decisions. We found that consumers are more likely to shy away from products with high arousal potential designs when making decisions for longer-term use (studies 1a and 1b). The reason for this is that consumers expect to be more irritated by high (vs. low) arousal designs over time (studies 1 and 2) and therefore predict faster satiation and less use from these designs (studies 2 and 3). Importantly, however, these predictions seem to be misguided (studies 4 and 5). In both a lab and a natural field setting, we show that contrary to forecasters' predictions, experiencers tire more quickly of low arousal designs than of high arousal designs. This leads to systematic errors in predicted utility: Forecasters overestimate satiation from high arousal designs and as a result underestimate their liking of high arousal designs over time.

Theoretical Contributions

By showing a new type of forecasting error, the present research contributes to the work on predicted and experienced utility in two significant ways. First, it identifies differences in predicted and experienced utility in a new and marketing relevant consumption domain: visual processing of aesthetics. Past research on predicted and experienced utility has examined a variety of hedonic experiences and sensations, ranging from more general measures of happiness to more specific sensory measures such auditory and gustatory liking. The present research is, to the best of our knowledge, the first to compare predicted and experienced utility from visual stimuli. Second, it examines how characteristics of (visual) stimuli differently influence predicted and experienced liking over time. Previous research has established that consumers mispredict liking for isolated sensory stimuli over time (e.g., eating yoghurt, listening to music), but has not compared different levels of these discreet stimuli (e.g., the intensity of a flavor or music). By systematically manipulating the level of arousal potential for the same sensory stimulus, the present research provides important insight into when and why errors in predicted liking are more or less likely to occur. This focus on how characteristics of sensory stimuli determine errors in predicted liking complements and extends related efforts in the affective forecasting literature, which have examined how characteristics of events (e.g., their probability or psychological distance) differently influence their emotional impact in prospect and experience (e.g., Buechel et al. 2014; Ebert and Meyvis 2014).

The current research also makes important contributions to the literature on aesthetics in consumer behavior. Some of the most basic work on aesthetics considers how perceptual processes influence evaluation and liking (e.g., how fluency determines ease of processing and liking). We show how two distinctive design elements – color and pattern – exhibit similarities in how they shape preferences and hedonic value because they similarly influence the arousal potential of the stimulus. It follows that the findings of this research may be extrapolated to other manipulations of visual arousal potential such as the magnitude of contrast or movement in visual stimuli. Furthermore, while literature on design and mere exposure has identified factors that influence adaptation and satiation to visual stimuli, this past literature has not systematically examined predicted liking. This is notable, given that most consumption and purchase decisions are based on predicted hedonic value. By showing that forecasted liking does not match the dynamics of experienced liking, the present work fills an important gap that allows us to better understand how arousal potential influences judgments and decisions made for future consumption.

On a broader level, the current research might provide an important step to the better understanding sensory/hedonic adaptation and sensitization. By showing how sensitization on one measure (e.g., increase in irritation) can lead to adaptation on a different measure (e.g., decrease in liking or general hedonic response), we show how the two can be interrelated and inform each other. This highlights the importance of clarity and specificity in terminology when talking about dynamics of sensory and hedonic adaptation. Further, while we have focused on visual stimuli, the findings might point to a more general overarching theory that can help predict the dynamics of anticipated (vs. experienced) hedonic value over time. Past research has identified isolated instances of sensitization (e.g., unpalatable gustatory stimuli, loud and high pitch noises, headaches, sexual arousal) and adaptation (e.g., highway noise, sauna visits, mild elective shocks), but consumers and researchers seem to have a relatively poor understanding of the boundary conditions that determine when and to what extent each occurs (see Snell et al. 1995; Snell and Gibbs 1995). Our research suggests that arousal potential of a stimulus might be an important determinant of predicted and experienced hedonic value over time. If so, then our results might generalize to other modalities (e.g., gustatory, auditory). In other words, consumers might expect to grow more irritated and satiate more quickly from intense sounds (e.g., a new type of techno) or flavors (e.g., an intense curry), when in reality, these might bring continued pleasure.

Implications for Consumers and Marketers

Apart from being of theoretical interest, the present research has important implications for how consumers make decisions about aesthetic elements. When decorating homes and buying clothes, consumers often choose pale colors and simple designs (i.e., low arousal design items), presumably because they believe that this will increase long-term satisfaction with the product aesthetics. Our research suggests that by using this strategy, consumers might be leaving happiness on the table. Contrary to people's intuition, our studies suggest that consumers satiate more quickly from low arousal designs than from high arousal designs. The implication for consumers is that when making long-term decisions, consumers can be bolder without jeopardizing hedonic value. Even when making relatively permanent decisions about design, consumers should go for the intense colors and intense patterns they like – they might yield more hedonic value over time. Of course, consumers may want to shift their strategy when making choices on behalf of others. When preparing to sell a house, for example, sellers should recognize that a potential buyer unfamiliar with this research might falsely shy away from bold patterns, decreasing the resale value. Indeed, real estate and color consultants advise homeowners to "start with earth tones" and to "get rid of anything that's kind of obnoxious" (Moore 2010; realestate.aol.com).

The current research has equally significant implications for marketers. When designing logos, stores, websites, or any other design aspect associated with a product, marketers should again embrace bold product designs. Not only will they capture attention, but they are likely free of the feared downsides such as satiation and irritation.

Limitations and Future Research

We observed our hypothesized effects over a variety of product category and designs. Nevertheless, it is important to keep in mind the possible boundary conditions of this effect. Specifically, the importance and nature of aesthetic considerations might vary across product categories, which could diminish and exacerbate our findings. Some product categories such as clothes, for example, require combinations of multiple products/items that may vary on aesthetics. For such products, consumers might find additional motivation in selecting low-arousal designs because it makes matching different designs easier. Other products might be more likely to be subject to aesthetic trends and fashion considerations, again providing additional motivation to select low arousal designs. It is plausible that consumers veer away from bright colors and busy patterns when purchasing for long-term use based on the presumption that such high arousal elements are more likely to go out of style over time than more mundane ones. To mitigate the issues of matching and trend, we selected a variety of different products, including ones that could be consumed in isolation (e.g., iPad cover) and ones for which fashion or trends were not highly relevant (e.g., name plates). Yet, it is important to recognize that our effects might be weaker in some product categories than others.

This discussion highlights the potential for future research on moderators in this area, such as for which product categories misprediction is more or less likely to occur. As already hinted at, another interesting avenue for future research would be to see whether our results hold for other sensory domains such as manipulations of auditory loudness or complexity of taste.

Overall, the results consistently show (misguided) expectations of satiation as a primary driver of the decreased preference for high arousal design for the long run. Importantly, as a high level of functional quality becomes standard across numerous product categories (Page and Herr 2002), more brands are looking to aesthetics as a point of differentiation. Therefore, this research is relevant to an increasingly diverse range of product purchases and usage experiences.

REFERENCES

Alba, Joseph and Elanor F. Williams (2013), “Pleasure Principles: Current Research on Hedonic Consumption,” *Journal of Consumer Psychology*, 23 (1), 2-18.

Aiken, Leona and Stephen G. West (1991), *Multiple Regression: Testing and Interpreting Interactions*, Newbury Park, CA: Sage.

Baker, Garry and Robert Franken (1999), "Effects of Stimulus Size, Brightness and Complexity upon EEG Desynchronization," *Psychonomic Science,* 7 (9), 289-90.

Berlyne, Daniel E. (1960), "*Conflict, Arousal, and Curiosity*." New York, NY: McGraw-Hill.

_______ (1968), "The Motivational Significance of Collative Variables," in *Theories of Cognitive Consistency: A Sourcebook,* Ed. Robert Abelson, Chicago, IL: Rand McNally, 275-88.

_______ (1970), "Novelty, Complexity, and Hedonic Value," *Perception & Psychophysics*, 8 (5), 279-86.

Berlyne, D. E. and P. McDonnell (1965), "Effects of Stimulus Complexity and Incongruity on Duration of EEG Desynchronization," *Electroencephalography and Clinical Neurophysiology,* 18 (2), 156-61.

Berlyne, Daniel E. and George H. Lawrence (1964), "Effects of Complexity and Incongruity Variables on GSR, Investigatory Behavior and Verbally Expressed Preference," *Journal of General Psychology*, 71 (1), 21-45.

Bornstein, Robert F. (1989), "Exposure and Affect: Overview and Meta-Analysis of Research, 1968–1987," *Psychological Bulletin,* 106 (2), 265 - 89.

Buechel, Eva C., Jiao Zhang, Carey K. Morewedge, and Joachim Vosgerau (2014), "More Intense Experiences, Less Intense Forecasts: Why People Overweight Probability

Specifications in Affective Forecasts," *Journal of Personality and Social Psychology,* 106 (1), 20-36.

Cantor, Gordon N. (1968), "Children's 'Like-Dislike' Ratings of Familiarized and Non-familiarized Visual Stimuli," *Journal of Experimental Child Psychology*, 6 (4), 651-57.

Coombs, Clyde H. and George S. Avrunin (1977), "Single-Peaked Functions and the Theory of Preference," *Psychological Review*, 84 (2), 216 - 30.

Cox, Dena and Anthony D. Cox (2002), "Beyond First Impressions: The Effects of Repeated Exposure on Consumer Liking of Visually Complex and Simple Product Designs." *Journal of the Academy of Marketing Science,* 30 (2), 119-30.

Ebert, Jane E.J and Tom Meyvis (2014), "Reading Fictional Stories and Winning Delayed Prizes: The Surprising Emotional Impact of Distant Events," *Journal of Consumer Research*, 41 (3), 794 – 809.

Fitzsimons, Gavan J. (2008), "Death to Dichotomizing," *Journal of Consumer Research*, 35 (1), 5-8.

Frederick, Shane and George Loewenstein (1999), ''Hedonic Adaptation,'' in *Well-being: The foundations of Hedonic Psychology*, Ed. Daniel Kahneman, Ed Diener, and Norbert Schwarz, New York, NY: Sage, 302-29.

Gilbert, Daniel T., Elizabeth C. Pinel, Timothy D. Wilson, Stephen J. Blumberg, and Thalia P. Wheatley (1998), "Immune Neglect: A Source of Durability Bias in Affective Forecasting," *Journal of Personality and Social Psychology*, 75 (3), 617-38.

Hayes, Andrew F. (2013), *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-based Approach*, New York, NY: Guilford Press.

Heiman, Julia R. (1977), "A Psychophysiological Exploration of Sexual Arousal Patterns in Females and Males," *Psychophysiology,* 14 (3), 266-74.

Hoegg, JoAndrea and Joseph. W. Alba (2008), "A Role for Aesthetics in Consumer Psychology," in *Handbook of Consumer Psychology*, Ed. Frank Kardes, Curtis Haugtvedt and Paul Herr, New York, NY: Psychology Press, 733-54.

Holtzschue, Linda (2011), *Understand Color: An Introduction For Designers,* 4th Edition, Hoboken, NJ: John Wiley & Sons, Inc.

Janiszewski, Chris and Tom Meyvis (2001), "Effects of Brand Logo Complexity, Repetition, and Spacing on Processing Fluency and Judgment," *Journal of Consumer Research,* 28 (1), 18-32.

Kahneman, Daniel and Jackie Snell (1992), "Predicting a Changing Taste: Do People Know What They Will Like?" *Journal of Behavioral Decision Making*, 5(3), 187-200.

Küller, Rikard, Byron Mikellides, and Jan Janssens (2009), "Color, Arousal, and Performance – A Comparison of Three Experiments," *Color Research & Application,* 34 (2), 141-52.

Laan, Ellen, Walter Everaerd, and Andrea Evers (1995), "Assessment of Female Sexual Arousal: Response Specificity and Construct Validity," *Psychophysiology*, 32 (5), 476-85.

Mehrabian, Albert and James A. Russell (1973) "A Measure of Arousal Seeking Tendency." *Environment and Behavior*, 5 (3), 315-33.

Moore, Christopher (2010), "Painting to Sell: What Color Homes Sell Best?" http://realestate.aol.com/blog/2010/07/16/painting-to-sell-what-color-homes-sell-best/

Nelson, Leif D. and Tom Meyvis (2008), "Interrupted Consumption: Disrupting Adaptation to Hedonic Experiences," *Journal of Marketing Research*, 45 (6), 654-64.

Nicki, Richard M. and Anthony Gale (1977) "EEG, Measures of Complexity, and Preference for Nonrepresentational Works of Art," *Perception*, 6, 281-86.

Page, Christine and Paul M. Herr (2002), "An Investigation of the Processes by which Product Design and Brand Strength Interact to Determine Initial Affect and Quality Judgments," *Journal of Consumer Psychology*, 12(2), 133-47.

Patrick, Vanessa A. and Laura A. Peracchio (2010), "Curating the JCP Special Issue on Aesthetics in Consumer Psychology: An Introduction to the Aesthetics Issues," *Journal of Consumer Psychology*, 20 (4), 393-97.

Postrel, Virginia (2003), *The substance of style*. New York, NY: Harper Collins.

Redden, Joseph P. (2008), "Reducing Satiation: The Role of Categorization Level," *Journal of Consumer Research*, 34 (5), 624-34.

Sieff, Elaine M., Robyn M. Dawes, and George Loewenstein (1999), "Anticipated versus Actual Reaction to HIV Test Results," *The American Journal of Psychology,* 112 (2), 297-311.

Simonson, Itamar (1990), "The effect of Purchase Quantity and Timing on Variety-Seeking Behavior," *Journal of Marketing Research*, 27 (2), 150-62.

Snell, Jackie and Brian J. Gibbs (1995) "Do Consumer Know What They Will Like?" *Advances in Consumer Research*, 22, 277-79.

Snell, Jackie, Brian J. Gibbs, and Carol Varey, (1995), "Intuitive Hedonics: Consumer Beliefs about the Dynamics of Liking," *Journal of Consumer Psychology*, 4 (1), 33-60.

Townsend, Claudia and Suzanne B. Shu (2010), "When and How Aesthetics Influences Financial Decisions," *Journal of Consumer Psychology*, 20 (4), 452-58.

Weinstein, Neil D. (1978), "Individual Differences in Reactions to Noise: A Longitudinal Study in a College Dormitory," *Journal of Applied Psychology,* 63 (4), 458 -66.

_______ (1982), "Community Noise Problems: Evidence against Adaptation," *Journal of Environmental Psychology,* 2 (2), 87-97.

Weller, Leonard and Randy Livingston (1988), “Effects of Color of Questionnaire on Emotional Responses,” *Journal of General Psychology*, 115 (4), 433-40.

Wright, Benjamin and Lee Rainwater (1962), “The Meanings of Color,” *Journal of General Psychology*, 67 (1), 89-99.

Wilson, Timothy D. and Daniel T. Gilbert (2003), "Affective Forecasting," *Advances in Experimental Social Psychology*, 35 (December), 345-411.

Wilson, Timothy D. and Daniel T. Gilbert (2005), "Affective Forecasting: Knowing What to Want," *Current Directions in Psychological Science*, 14 (3), 131-34.

Wundt, Wilhelm M. (1874), “*Grundzüge der Physiologischen Psychologie [Translation: Principles of Physiological Psychology],”* Leipzig: Engelmann.

Yamamoto, Mel and David R. Lambert (1994), "The Impact of Product Aesthetics on the Evaluation of Industrial Products," *Journal of Product Innovation Management,* 11 (4), 309-24.

Zajonc, Robert B. (1968). "Attitudinal Effects of Mere Exposure," *Journal of Personality and Social Psychology,* 9 (2), 1-27.

Zajonc, Robert B., Philip Shaver, Carol Tavris, and David Van Kreveld (1972), "Exposure, Satiation, and Stimulus Discriminability," *Journal of Personality and Social Psychology*, 21 (3), 270 - 80.

Zielińki, Piotr (2016), “An Arousal Effect of Colors Saturation: A Study of Self-Reported and Electrodermal Responses,” *Journal of Psychophysiology*, 30 (1), 9-16.

Zhao, Xinshu, John G. Lynch, and Qimei Chen (2010), "Reconsidering Baron and Kenny: Myths and Truths about Mediation Analysis," *Journal of Consumer Research,* 37 (2), 197-206.

APPENDIX 1

Stimuli used

Low Arousal **High Arousal**

Study 1a: Paper Cups

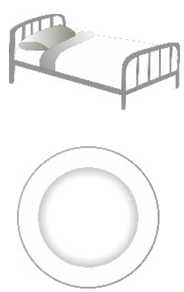
Participants saw the low arousal design and a randomly selected high arousal design.

Study 1b: iPad Covers

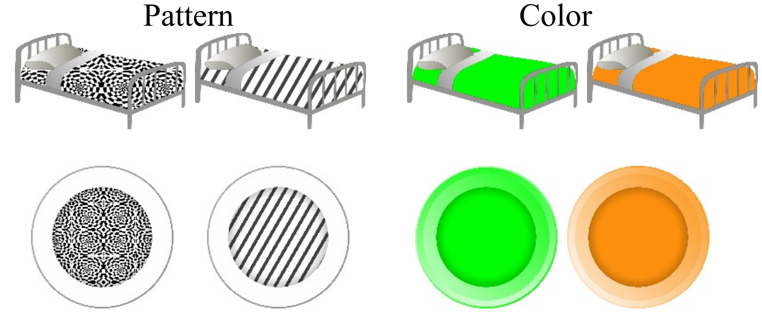
Participants saw a randomly selected pair – one low arousal, one high arousal design.

Study 2: Beds and Plates

Participants saw both product categories (order randomized) with a randomly selected product design.



Pattern Color

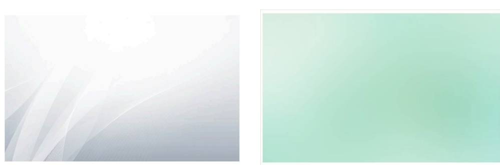


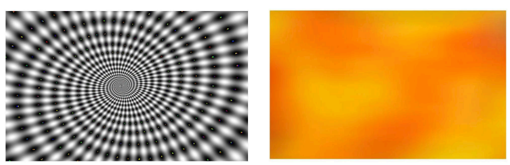
<u>Study 3: Workout Shirts</u>

Participants saw one randomly selected shirt design.
Images display male shirts; females were shown a more fitted shape.

<u>Study 4: Screen Backgrounds</u>

Participants saw one randomly selected screen background.





<u>Study 5: Name Cards</u>

Participants were randomly assigned to one name card.