Summer 2020

Testing for Differential Attention to Features in Evaluative Conditioning

Christine E. Weber

TESTING FOR DIFFERENTIAL ATTENTION TO FEATURES IN EVALUATIVE CONDITIONING

by

Christine E. Weber

Bachelor of Science
Eastern Connecticut State University, 2015

Master of Arts
University of South Carolina, 2018

Submitted in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy in

Experimental Psychology

College of Arts and Sciences

University of South Carolina

2020

Accepted by:

Douglas H. Wedell, Major Professor

Amit Almor, Committee Member

Abhijit Guha, Committee Member

Svetlana V. Shinkareva, Committee Member

Cheryl L. Addy, Vice Provost and Dean of the Graduate School

ACKNOWLEDGEMENTS

I would like to express my immense gratitude to my advisor, Dr. Douglas H. Wedell, for his guidance, mentorship, and advice throughout my graduate career. I would also like to thank my other committee members, Dr. Svetlana V. Shinkareva, Dr. Amit Almor, and Dr. Abhijit Guha, who have given me the benefit of their knowledge, experience, and support. I would like to thank Chuanji Gao, William Hayes, and the other members of my lab for their contributions to study design, data collection, and analysis.

Thank you always to my family, my fiancé, Jeffrey Vahlstrom, and my friend, Katrina Macher, for their love and support.

ABSTRACT

One eye tracking and one behavioral experiment examined the possible roles of contingency awareness and attention to stimulus features in evaluative conditioning. These experiments tested whether evaluative conditioning altered the saliency of positive and negative features in consumer products (Study 1) and attitudinal responses to ambivalent pictures (Study 2). Based on the conceptual categorization model, pairing of ambivalent conditioned stimuli with liked or disliked unconditioned stimuli was predicted to result in enhanced attention to affectively congruent features. Study 1 tested this prediction by recording eye movements to determine how attention to features and responses to ambivalent stimuli were altered as a result of conditioning. Counter to the predictions of the conceptual categorization model, pairing of ambivalent products with liked or disliked music did not result in greater attention to affectively congruent features as measured by looking time and frequency. Overall, target features that were inconsistent in valence with the unconditioned stimulus were looked at longer and more frequently. Studies 1 and 2 tested whether conditioning effects were consistent with predictions from explicit process models that contingency awareness should positively correlate with conditioning effects. Study 2 was more strongly powered to detect this relation and showed that contingency awareness was a significant predictor of conditioning effects, consistent with an explicit mechanism in evaluative conditioning. Results from these studies demonstrate evaluative conditioning of ambivalent stimuli that is not likely due to enhanced attention to affectively congruent features and that is consistent with explicit processing accounts.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Using evaluative conditioning, preferences for neutral targets such as consumer products can be shifted by pairing them with affectively charged stimuli (De Houwer, Thomas & Baeyens, 2001; Hofmann, De Houwer, Perugini, Baeyens, & Crombez, 2010; Kardes, Posavac, & Cronley, 2004; Martin & Levey, 1978). Two studies investigated the potential roles of differential attention to affectively congruent features and contingency awareness in evaluative conditioning of ambivalent stimuli. In particular, to the degree that conditioned stimuli may be ambivalent in nature, pairing them with a liked or disliked unconditioned stimulus is hypothesized to lead to enhanced attention to affectively congruent features. Ambivalent stimuli may be defined as having both positive and negative attributes (Priester & Petty, 1996). The experiments here tested this hypothesis using behavioral and eye tracking methods and included measures of contingency awareness to examine whether the conditioning effects depended on awareness of the pairings. Before delving into the details of this account, I first review the literature on models of evaluative conditioning and the role of contingency awareness.

1.1 MODELS OF EVALUATIVE CONDITIONING

In evaluative conditioning, affective and attitudinal responses to target stimuli can become either positive or negative by pairing the targets with positive or negative unconditioned stimuli. There is a comprehensive literature demonstrating evaluative conditioning effects with a wide variety of stimulus types, conditioning procedures, and dependent measures. Although the effects of evaluative conditioning are therefore well-established and generalizable to many contexts, there are several competing models to explain the specific mechanisms underlying these effects. These models propose the effects are the result of implicit or explicit processing, or perhaps both. The empirical support for underlying mechanisms is mixed, indicating multiple mechanisms may be at work or differentially applicable across tasks and stimulus conditions.

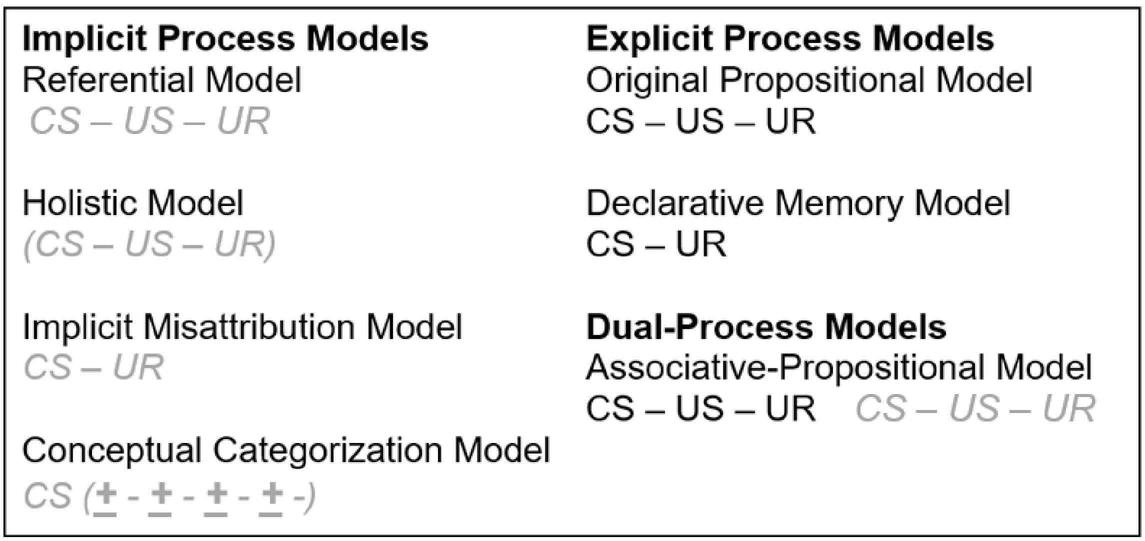

Figure 1.1 Models of evaluative conditioning. Theoretical learning representations of implicit, explicit, and dual-process models of evaluative conditioning. Gray italicized text indicates implicit representations. CS = conditioned stimulus, US = unconditioned stimulus, UR = unconditioned response

Most models of evaluative conditioning can be categorized into two classes: implicit process models and explicit process models (Figure 1). The primary distinction between implicit and explicit process models is that implicit process models propose evaluative conditioning is due to automatic processing operating below conscious awareness, and explicit process models propose that evaluative conditioning is the result of non-automatic, deliberate conscious processes (Hofmann et al., 2010). The implicit process models therefore do not posit the necessary involvement of conscious knowledge about the pairings or contingencies between stimuli, while the explicit process models generally require some degree of awareness about stimuli relations. There are several implicit process models which have been proposed in the past three decades, including the referential model (Baeyens, Eelen, Crombez, & Van den Bergh, 1992), the holistic model (Martin & Levey, 1994), the implicit misattribution model (Jones, Fazio, & Olson, 2009), and the conceptual categorization model (Field & Davey, 1999). There have also been various instantiations of explicit process models, including the propositional model (Baeyens, Field, & De Houwer, 2005) and the declarative memory model (Gast, 2018). Although most of the implicit process models are often called associative models, it is important to recognize that the explicit models do not argue against the involvement of associations in evaluative conditioning. There is also one model which combines features of both types of processing: the associative-propositional evaluation (APE) model (Gawronski & Bodenhausen, 2006).

1.1.1 IMPLICIT PROCESS MODELS

In general, the implicit process models describe evaluative conditioning effects as the result of automatic associations between the conditioned stimulus and either the unconditioned stimulus, the unconditioned response, or both. The referential, holistic, and implicit misattribution models differ in the specific mechanism by which evaluative conditioning is proposed to occur but share many predictions, whereas the conceptual categorization model is a non-associative account with unique predictions.

The referential, holistic, and implicit misattribution models of evaluative conditioning are implicit associative models. The referential model posits that the response to the conditioned stimulus is altered due to activation of a representation of the unconditioned stimulus upon subsequent presentation of the conditioned stimulus (Baeyens et al., 1992). In other words, an association is formed between the unconditioned stimulus and the conditioned stimulus. The holistic model proposes that the response to the conditioned stimulus is altered due to the creation of a “holistic” representation that includes three elements: the unconditioned stimulus, the conditioned stimulus, and the response (which is shared) (De Houwer et al., 2001; Hofmann et al., 2010; Martin & Levey; 1994). In the holistic model, presentation of the conditioned stimulus results in activation of this combined representation, therefore producing the response associated with the unconditioned and conditioned stimuli (Martin & Levey, 1994). The implicit misattribution model proposes that the relevant association is between the conditioned stimulus and the response to the unconditioned stimulus, rather than the unconditioned stimulus itself (Jones et al., 2009; Hofmann et al., 2010). Furthermore, the model proposes a specific explanation for formation of this association – namely, that the response to the unconditioned stimulus is mistakenly attributed to the conditioned stimulus, and that this occurs below the level of conscious awareness (Jones et al., 2009). For the purposes of this dissertation, the predictions of these three implicit process models are largely overlapping and are not differentiated from one another in these experiments.

The conceptual categorization model (Field & Davey, 1999) proposes a rather unique mechanism which distinguishes it from the other implicit process models, as well as from the explicit process models. Unlike these other models, which all acknowledge the role of either implicit or explicit links between the unconditioned and conditioned stimuli, the conceptual categorization model is a non-associative account. This model proposes that the response to the conditioned stimulus is altered due to a recategorization of the conditioned stimulus from a neutral category to a liked/positive or disliked/negative category (Davey, 1994; Field & Davey, 1999). Additionally, this recategorization is posited to result from increase saliency of shared or similar features between the unconditioned and conditioned stimulus (Field & Davey, 1999; Hofmann et al., 2010). The pairing of the two stimuli highlights the similar features of the conditioned stimulus, resulting in increased attention to those features and a change in its evaluative category. This model proposes that evaluative conditioning is a form of conceptual rather than associative learning, more similar to learning what features exemplify a category than to Pavlovian conditioning. For the current experiments, an expanded version of this model was used that predicted attention is increased for shared affective features, rather than shared perceptual features, thus underlying evaluative conditioning effects. It was hypothesized that ambivalent stimuli containing both liked and disliked features would become more liked or more disliked by shifting attention to affectively congruent features via conditioning in Study 1. Thus, Study 1 tested a modified version of the conceptual categorization model based on shared affective features rather than shared perceptual features.

1.1.2 EXPLICIT PROCESS AND DUAL-PROCESS MODELS

Explicit process models argue that evaluative conditioning is the result of non-automatic, conscious processes. These models are generally more recent than the more classical implicit associative models.

The original propositional model was first proposed by Baeyens and colleagues (2005) and was expanded in their later work (Bar-Anan, De Houwer, & Nosek, 2010; De Houwer, 2006; De Houwer, 2007; Mitchell, De Houwer, & Lovibond, 2009). This model proposes that the response to the conditioned stimulus is altered due to the use of conscious, explicit knowledge about the relationship between the unconditioned and conditioned stimulus. Its main proposition is that evaluative conditioning relies on 1) the creation of a propositional statement about the relation or contingency between the two stimuli, and 2) the use of this knowledge in evaluating the conditioned stimulus, based on a belief that this propositional statement is valid (Baeyens et al. 2005; Hofmann et al., 2010). Importantly, this view still necessarily involves a link between the two stimuli, however, the link is described as explicit and non-automatic. The evidence supporting the propositional model comes primarily from studies showing that conditioning effects are dependent on conscious knowledge of the relationship between the unconditioned and conditioned stimuli (e.g. Blask, Walther, Halbeisen, & Weil, 2012; Field & Moore, 2005; Stahl, Unkelbach, & Corneille, 2009).

The declarative memory model is the most recently proposed account of evaluative conditioning. Like the original propositional model, the declarative memory model ascribes evaluative conditioning effects to the use of conscious, explicit knowledge about the relation between the conditioned and unconditioned stimuli. However, it is more specific in its characterization of the mechanism involved. As described in the original article (Gast, 2018), the declarative memory model describes evaluative conditioning effects as the result of four steps. First, presentation of the unconditioned and conditioned stimuli results in an association in memory between the conditioned stimulus and evaluative information about the unconditioned stimulus; second, this association still exists at testing, third; this association is consciously accessed at testing; and fourth, this association is used in forming a response at testing (Gast, 2018). This model is supported by evidence showing that participants' conscious knowledge of stimulus associations, and the use of that knowledge at test, predict the strength of conditioning effects (Gast, 2018).

The APE model is a dual-process model that describes the involvement of both implicit associations and explicit propositions in evaluative conditioning. Unlike the previously discussed accounts, this model is based on a theoretical distinction between implicit and explicit attitudes. It argues that implicit attitudes are primarily the result of automatic associations, whereas explicit attitudes are primarily the result of conscious proposition evaluations (Gawronski & Bodenhausen, 2006; Gawronski & Bodenhausen, 2011). However, these two processes are not assumed to be independent, rather, each should affect the other, with automatic associations often the basis of propositional statements, and propositional statements having the capacity to alter activation related to automatic associations (Gawronski & Bodenhausen, 2006). Therefore, this model proposes that the response to the conditioned stimulus is altered due to the interaction of an implicit associative process and a conscious propositional evaluation process. This dual-process account is supported by the mixed evidence for the effect of awareness of stimuli relations in evaluative conditioning, with some studies showing effects depend on conscious knowledge (Blask et al., 2012; Field & Moore, 2005; Stahl et al., 2009) and others showing effects with no evidence of conscious knowledge (Baeyens et al., 1992; Gawronski & Mitchell, 2014; Hütter & Sweldens, 2013).

1.2 FEATURE SALIENCY AND AMBIVALENT STIMULI

Unconditioned stimuli and target stimuli may possess shared features. While the implicit misattribution model proposes that evaluative conditioning effects are more likely to occur when the conditioned and unconditioned stimulus share perceptual features, this prediction has not been well-supported. The number of studies investigating the moderating role of stimulus similarity on evaluative conditioning effects is limited (Baeyens, Eelen, Van den Bergh, & Crombez, 1989; Jones et al., 2009). One study found that spatial proximity of the stimuli, gaze shifts between the stimuli, and increased salience of the conditioned stimuli increased the effect (Jones et al., 2009). However, a study investigating the specific effect of perceptual similarity found no difference in conditioning when stimuli were similar or dissimilar (Baeyens et al., 1989). Pairing similar conditioned and unconditioned visual stimuli (two faces or two art pieces) did not produce greater effects than pairing dissimilar stimuli (face and art). It can be assumed that stimuli within a modality tend to share more perceptual features than stimuli from two different modalities. Effect sizes in evaluative conditioning are similar regardless of whether the two stimuli share a common modality, strongly implying that perceptual similarity is not a significant moderator of conditioning effects (Hofmann et al., 2010).

The presence of shared affective features, on the other hand, may underlie recategorization of the target stimuli in line with the predictions of the conceptual categorization model. Stimuli that are not perceived as wholly positive or negative may still contain positive and negative features. For example, a photograph may contain both liked elements (a puppy or a favorite celebrity) and disliked elements (a spider or a displeasing color). One recent study found that some pictures from the International Affective Picture System (IAPS) categorized as neutral were actually ambivalent, rated as simultaneously positive and negative (Schneider, Veenstra, van Harreveld, Schwarz, & Koole, 2016). Although the potential moderating effect of shared affective features on evaluative conditioning does not link well with the perceptual similarity assertion of the implicit misattribution model, this may be a mechanism underlying some evaluative conditioning effects in line with the predictions of the conceptual categorization model.

Ambivalent stimuli with known positive and negative features can serve as ideal targets in testing the conceptual categorization model. To my knowledge, this is the first study demonstrating evaluative conditioning on stimuli designed to be ambivalent. If conditioning is observed, this would expand the generality of evaluative conditioning to these stimuli. If conditioning increases the saliency of affectively consistent features and shifts in visual attention are related to the magnitude of evaluative conditioning effects, then this would support the conceptual categorization model. If conditioning does not increase the saliency of affectively consistent features during conditioning, but the effects of conditioning remain, then this affective similarity-based version of the model would be unlikely.

1.3 CONTINGENCY AWARENESS

In addition to investigating the role of attention to shared affective features that test predictions unique to the conceptual categorization model, , the role of explicit processing was also considered. The primary prediction on which the implicit associative and propositional models differ is in the role of contingency awareness. Contingency awareness involves conscious knowledge of the relationship between the unconditioned and conditioned stimuli. Implicit process models posit that evaluative conditioning effects should be independent of contingency awareness, while explicit process models require contingency awareness. The dual-process APE model allows for circumstances in which contingency awareness may or may not modulate conditioning effects.

1.3.1 MEASURING CONTINGENCY AWARENESS

Contingency awareness can be measured in several ways. Participants may be able to report that a specific conditioned stimulus was always paired with an unconditioned stimulus of a specific valence (such as positive or negative). They may further be able to report with which specific unconditioned stimulus an individual conditioned stimulus was paired. In addition, participants can rate their confidence for each of these responses (e.g. Baeyens et al., 1992). The degree to which one can identify the unconditioned stimulus valence or the specific exemplar is an indicator of contingency awareness. Most evaluative conditioning studies that have measured awareness did so by presenting some or all of the unconditioned stimuli and asking participants to indicate which one each conditioned stimulus was paired with (Baeyens et al., 1992; Gawronski & Mitchell, 2014; Stahl et al., 2009; Walther & Nagengast, 2006), or presenting two stimuli and asking if they were paired together (Olson & Fazio, 2006). Others have asked participants to report the valence of the unconditioned stimulus in the pairing (Hütter et al., 2012; Pleyers, Corneille, Luminet, & Yzerbyt, 2007; Stahl & Unkelbach, 2009; Stahl et al., 2009). Open-ended questions assessing awareness have also been used (Olson & Fazio, 2006; Pleyers et al., 2007; Stuart, Shimp, & Engle, 1987; Walther, 2002).

1.3.2 EVIDENCE FOR CONTINGENCY AWARENESS

Many studies have tested the effects of contingency awareness in evaluative conditioning. Some studies have found evaluative conditioning effects are dependent on contingency awareness, with reduced or no effects without awareness (Blask et al., 2012; Field & Moore, 2005; Gast, De Houwer, & De Schryver, 2012; Stahl et al., 2009). Other studies have shown that conditioning effects remain without contingency awareness, although they may be reduced in some cases (Baeyens, Eelen, & Van den Bergh, 1990; Baeyens et al., 1992; Gawronski & Mitchell, 2014; Hütter & Sweldens, 2013; Hütter, Sweldens, Stahl, Unkelbach, & Klauer, 2012; Olson & Fazio, 2006; Pleyers et al., 2007). Some studies have even shown the opposite result of stronger effects without awareness (Walther, 2002).

As there are often considerable differences in the procedure and operationalization of evaluative conditioning experiments, some of this theoretical disagreement may be attributable to the specific paradigms used in different cases. For example, one study found conditioning effects without awareness only when the conditioned and unconditioned stimuli were paired simultaneously rather than sequentially (Hütter & Sweldens, 2013). Several other studies using simultaneous presentations have also shown effects without contingency awareness (Hütter et al., 2012; Olson & Fazio, 2006; Pleyers et al., 2007), although this finding has not been consistent in every such study (Blask et al., 2012; Gast et al., 2012). Despite the mixed effects associated with contingency awareness, it is useful to measure it to examine the degree to which the observed effects may be due to implicit or explicit processes.

1.4 AIMS OF THE CURRENT STUDIES

The primary aims of this research were two-fold. First, I tested whether evaluative conditioning effects were related to changes in attention during conditioning, whereby conditioning alter the saliency of positive and negative features in ambivalent stimuli as measured by eye fixation data. Specifically, I tested whether target features that were consistent in valence with the unconditioned stimuli were given more visual attention measured via looking time, as would be predicted by the conceptual categorization model.

Second, I tested to what extent these effects correlated with measures of contingency awareness to test between implicit and explicit mechanisms in evaluative conditioning. Together, the results of this set of studies help clarify the role of attention and awareness in evaluative conditioning of ambivalent target stimuli.

If evaluative conditioning effects are unrelated to contingency awareness, then this supports the role of implicit processing in evaluative conditioning. Furthermore, if conditioning alters visual attention to positive and negative features of ambivalent stimuli in a congruent manner, this lends support to a specific implicit process account, the conceptual-categorization model. If conditioning does not alter visual attention as predicted, this suggests that the representation formed during conditioning may be more holistic (Martin & Levey, 1994) and could be either associatively or propositionally based, with the relation to contingency awareness distinguishing between these possibilities. If evaluative conditioning effects are positively related to continency awareness, then this supports the role of contingency awareness and explicit processing in evaluative conditioning of ambiguous stimuli.

CHAPTER 2

STUDY 1 - EYE TRACKING WITH CONSUMER PRODUCTS

Study 1 was an eye tracking study testing how visual attention to positive and negative features of stimuli during conditioning may be altered by the valence of the unconditioned stimulus. Pictures of consumer products were paired with liked and disliked music to shift preferences. The product pictures were presented with positive and negative attributes in the form of four- and five-star or one- and two-star dimensional ratings, respectively.

This experiment followed directly from previous research investigating conditioning of consumer preferences using eye tracking (Weber, 2018). In a series of behavioral and eye tracking experiments, we tested evaluative conditioning effects by pairing pictures of consumer products with short positive and negative classical music clips. In two eye tracking studies, we found significant effects of preference on looking measures for the product pictures, with participants looking significantly longer and more often at the products they preferred during test. These results supported the correspondence between looking time and preference (Wedell & Senter, 1997).

Using significance tests on each individual's choice proportions for products paired with positive and negative classical music stimuli, 53 participants (42%) across the three studies chose products paired with positive music significantly more often. However, a large group of participants (24%) showed significant effects in the opposite direction. In a different set of behavioral studies, we had shown that these positive and negative music clips produce highly reliable and strong effects when testing effects on internal affective states (Weber, , 2020), but effects appear to be more variable across individuals when testing effects on preference. Therefore, we created a new stimulus set consisting of strongly liked popular music and strongly disliked unpopular music. Based on a more recent study using these new music stimuli, we found that this music produced stronger effects in the predicted direction and fewer reversed effects (described in more detail in section 2.1.2). Therefore, those liked and disliked music clips were used in Study 1 rather than positive and negative classical music clips, thereby enhancing effect sizes at both the individual and group level.

One of the previous experiments presented the product pictures along with the star ratings on various attributes in the test phase. Using eye tracking, we tested whether affectively congruent attributes would receive more attention measured via total looking time and total number of looks (Weber, 2018). Counter to our hypothesis, we found that participants did not spend more time looking at the congruent information. However, the conceptual categorization model would predict that the unconditioned stimuli should influence the saliency of congruent features in the conditioned stimuli during the conditioning phase, when the conditioned and unconditioned stimuli are presented simultaneously. Therefore, Study 1 presented these star attributes during conditioning and testing (rather than only at test) and thereby constitutes a much stronger test of the model's predictions.

This study tests how evaluative conditioning alters responses to visual stimuli made ambivalent by presenting them with positive and negative attributes, and how those responses are related to changes in visual attention to the attributes. In addition, this study tests whether contingency awareness moderates these effects. In line with the pattern of previous findings, particularly those studies using simultaneous presentation, I hypothesized that contingency awareness would correlate with the magnitude of evaluative conditioning effects, with greater awareness predicting a larger effect. If evaluative conditioning effects are independent from contingency awareness measures, then this supports implicit process models. If evaluative conditioning effects are related to contingency awareness, then this supports explicit or dual-process models.

2.1 METHOD

2.1.1 PARTICIPANTS

We recruited 25 University of South Carolina undergraduates for this experiment using the online SONA psychology participant pool. The original sample size goal was 50 participants, but only 25 were recruited due to the closing of the experimental facilities as a result of COVID-19. Participants were awarded extra credit in psychology courses in return for their participation and gave their written informed consent in accordance with the University of South Carolina Institutional Review Board.

The appropriate sample size was determined using a power analysis based on the average effect size (0.79) from three previous evaluative conditioning experiments using similar stimuli (Weber, Shinkareva, Kim, Gao, & Wedell, 2020) found that a sample size of 25 provided a power greater than .99 to detect the effect of conditioning. Previous studies have found significant effects of contingency awareness on evaluative conditioning with sample sizes of $n = 16$ (Stahl et al., 2009), $n = 30$ (Hütter & Sweldens, 2013), and $n = 36$ (Walther & Nagengast, 2006). The original sample size goal was designed to provide a more powerful test of the correlation with contingency awareness and better observe individual difference patterns.

2.1.2 MATERIALS

The unconditioned stimuli were liked and disliked music clips. These were short four-second music clips with vocalization collected in a series of norming studies. The liked music was clipped from top Billboard charts from the previous decade and was consistently rated as highly liked (M = 7.37 on a 9-point scale, range = 7.17 – 7.63) by a sample of University of South Carolina students (n = 47). The disliked music was heavy metal music clipped from songs listed on the iTunes Top Heavy Metal chart and was consistently rated as highly disliked (M = 1.67, range = 1.52 – 2.00). This music was selected on the basis of prior research conducted both by our own lab and by others. Previous literature has shown general dislike for heavy metal (Perham & Sykora, 2012). In a study conducted in our lab, this set of liked and disliked music was shown to produce more consistent evaluative conditioning effects than our previously used positive and negative music, with a significant effect at the group level and a large portion of participants (11 of 24) showing significant effects at the individual level in the anticipated direction and only a few showing significant reverse effects (5 of 24).

The conditioned stimuli were pictures of consumer products used in previous evaluative conditioning experiments in our lab (Weber, 2018). There were 60 pictures of products obtained from the “Bed, Bath, and Beyond” company website, with 12 pictures in each of five categories: dinnerware sets, floor lamps, portable speakers, throw pillows, and water bottles. Products within a category differed primarily by color and pattern.

The conditioned stimuli were presented with star ratings of attributes (Figure 2.1). Each product was presented with ratings on four attributes: durability, ease of use, reliability, and value, in the form of stars ranging from one star to five stars. For each product, two of these attributes were positive (four or five stars) and two attributes were negative (one or two stars). Sets of attributes were randomly paired with products, and ratings for a given product remained constant across each presentation. Across the four attributes, there were always 12 stars in total for each product, so that attribute values between the liked and disliked conditions were well-matched.

Figure 2.1 Example of a consumer product picture in Study 1. Products were presented with two positive (four- or five-star) and two negative (one- or two-star) attributes.

2.1.3 PROCEDURE

The procedure included five phases: pre-induction ratings, induction, paired choice, post-induction rating, and contingency awareness measure (Table 2.1). Trial order within each phase was fully randomized, except that induction and paired choice were split into two halves, with this separation undetectable for participants. There were no time limits on responses. All phases of the study were conducted using an SR Research Eyelink 1000 eye tracker device, and the experimental program was run in Experiment Builder. Participants viewed stimuli on a 19” CRT monitor using a resolution of 1024 × 768 seated approximately 24” from the screen. Product stimuli were displayed at 478 × 478 pixels (including a minimal white border) on a white background, subtending approximately 15.72° by 15.72° of visual angle. Participants placed their heads in a frame with chin and forehead rests in order to minimize head movement. Only the right eye was tracked.

Table 2.1 Materials and procedures used in Study 1.

Induction	Test	Contingency Awareness Measure
Trials: 160 (2 music conditions × 5 product categories × 2 exemplars × 8 repetitions) **Task:** Answer yes/no perceptual questions Preceded by **pre-induction ratings** (identical to post-induction ratings)	**Paired Choice** **Trials:** 160 (5 product categories × 4 combinations × 8 repetitions) **Task:** Choose the product you prefer **Post-Induction Ratings** **Trials:** 60 (5 product categories × 12 exemplars) **Task:** Rate liking	**Trials:** 20 (2 music conditions × 5 product categories × 2 exemplars) **Task:** Choose the product that was presented with the music clip

During the pre-induction phase, participants indicated how much they liked each product on a Likert-type scale, from 1 (*Dislike Very Much)* to 9 (*Like Very Much).* These ratings were used to select the products that were carried forward into the induction phase to be paired with music. Two pairs of products in each category were selected, by choosing pairs of products that were closely matched in liking and not extremely liked or disliked, with ratings close to 5 on the 9-point scale. The most neutral, closely matched products were selected from the full set to maximize the ability of the unconditioned music to influence the evaluations of the products.

In the induction phase, these twenty selected products (four products in each of five categories) were paired with music, so that within each pair of similarly rated products, one was paired with liked music while the other was paired with disliked music. Each picture was displayed for 4 s with its star attribute ratings while a music clip played simultaneously. Each product-music pairing was presented 8 times for a total of 160 trials. Following each presentation, participants were asked to answer a simple yes or no question about a perceptual feature of the product picture. They were asked to indicate whether or not they saw a particular feature: red, blue, green, yellow, ceramic, metal, fabric, or curves. Using an active task during induction helped maintain participants' attention to the stimuli presentations. These questions were used in previous experiments and yielded strong evaluative conditioning effects (Weber, 2018).

In the paired choice phase, participants were presented with each product pair and were asked to select the product they preferred by clicking on it using the mouse. Each pair was presented eight times, with each product presented on the left side of the screen for half of the trials. Each product was presented with its star attribute ratings from the previous phase. There were two products paired with liked music and two products paired with disliked music within each product category, and therefore four cross-valence pairs per category, resulting in 32 trials per category and 160 trials total.

The post-induction rating phase was identical to the pre-induction rating phase. Participants rated all 60 products again, including those that had and had not been paired with music in the induction phase.

In the contingency awareness measure, participants heard a music clip and saw all 20 products used in the induction phase presented on the screen at the same time. They were asked to select the product picture that they thought was shown when they heard that music clip before.

In each phase, the product images were the main areas of interest for eye tracking data collection. In the induction and paired choice phases, the attribute star ratings were additional areas of interest.

2.2 RESULTS

2.2.1 BEHAVIORAL DATA

A one-sample *t*-test was conducted to test the effect of evaluative conditioning on paired choice selections, testing the hypothesis that products presented with liked music would be chosen more often than products presented with disliked music. This test determined if the proportion of times products paired with liked music were chosen was significantly different from chance, with proportions transformed using an arcsine square root transformation. At the group level, the effect of conditioning on paired choice did not reach significance, $t(24) = 1.998$, $p = .057$. Products previously paired with liked music were selected 54.45% of the time.

Binomial tests were conducted to evaluative significance at the individual participant level. Nine participants chose products paired with liked music significantly more often; four participants chose products paired with disliked music significantly more often, and 12 showed no significant effect.

A repeated-measures ANOVA was conducted to test the effect of evaluative conditioning on product ratings, testing the hypothesis that products presented with liked music would be rated more highly than products presented with disliked music. A two-way repeated-measures ANOVA with factors music condition (liked or disliked) × product category was conducted on the change in ratings from pre-induction to post-induction across product categories. The main effect of valence did not reach significance, $F(1,24) = 0.685$, $p = .416$. The change in ratings did not differ between products paired with liked music ($M = 0.340$, $SD = 1.301$) and products paired with disliked music ($M = 0.200$, $SD = 1.253$). The main effect of product category was significant, $F(4,96) = 2.656$, $p = .038$, while the valence × product category interaction was not significant, $F(4,96) = 1.446$, $p = .225$ (although these effects are not of primary interest).

Follow-up analyses tested whether these effects were moderated by contingency awareness, testing the hypothesis that contingency awareness would predict evaluative conditioning effects in both test phases, with greater awareness predicting larger effects. A linear regression analysis was used to test whether the degree of an individual's contingency awareness (scored as number correct out of 20) predicted evaluative conditioning effects in the paired choice phase (scored as proportion of times products paired with liked music were chosen). Contingency awareness was not a significant predictor, $R^2 = .003$, $F(1,23) = 0.059$, $p = .810$, standardized $\beta = 0.051$. A parallel regression analysis using contingency awareness scored as number of trials with a product of the correct valence chosen found similar results.

Another linear regression analysis was used to test whether contingency awareness predicted evaluative conditioning effects in the post-induction rating phase (scored by taking the mean difference in ratings between products paired with liked versus disliked music). Contingency awareness was not a significant predictor, $R^2 = .067$, $F(1,23) = 1.654$,

p = .211, standardized β = 0.259. A parallel regression analysis using contingency awareness scored as number of trials with a product of the correct valence chosen found similar results.

Parallel regression analyses conducted on the subset of participants who showed a significant behavioral effect in paired choice were similarly nonsignificant, with the exception of the regression predicting rating change from contingency awareness calculated by matching valence, where contingency awareness was a significant predictor, R^2 = .258, F(1,11) = 5.179, p = .044, standardized β = 0.566.

2.2.2 EYE TRACKING DATA

For the following analyses on eye tracking data, analyses were conducted only with participants who had data in each cell of the design. Some missing cells occurred when there were no looks to specific interest areas. For this reason, the number of participants varies by analysis. In addition, fixations less than 50 ms were removed in all analyses following standard practice (Keating, 2014).

Repeated-measures ANOVAs were conducted on eye movement measures during the induction phase, testing the hypothesis that during conditioning, total number of fixations, total fixation durations, and first looks would be greater for affectively congruent attributes. A music valence (liked or disliked) × attribute valence (positive or negative) repeated-measures ANOVA was conducted on total number of fixations on the four attribute interest areas during induction (Figure 2.2, Panel A). There was a significant music valence × attribute valence interaction, F(1,24) = 4.719, p = .040, and a significant main effect of attribute valence, F(1,24) = 6.322, p = .019. The main effect of music valence was not significant, F(1,24) = 0.112, p = .740. For pictures presented with liked music, negative attributes ($M = 0.704$) were looked at significantly more often than positive attributes ($M = 0.410$). For pictures presented with disliked music, negative attributes ($M = 0.622$) were also looked at significantly more often than positive attributes ($M = 0.530$), although this difference was smaller as reflected in the significant music valence × attribute valence interaction (Table 2.2).

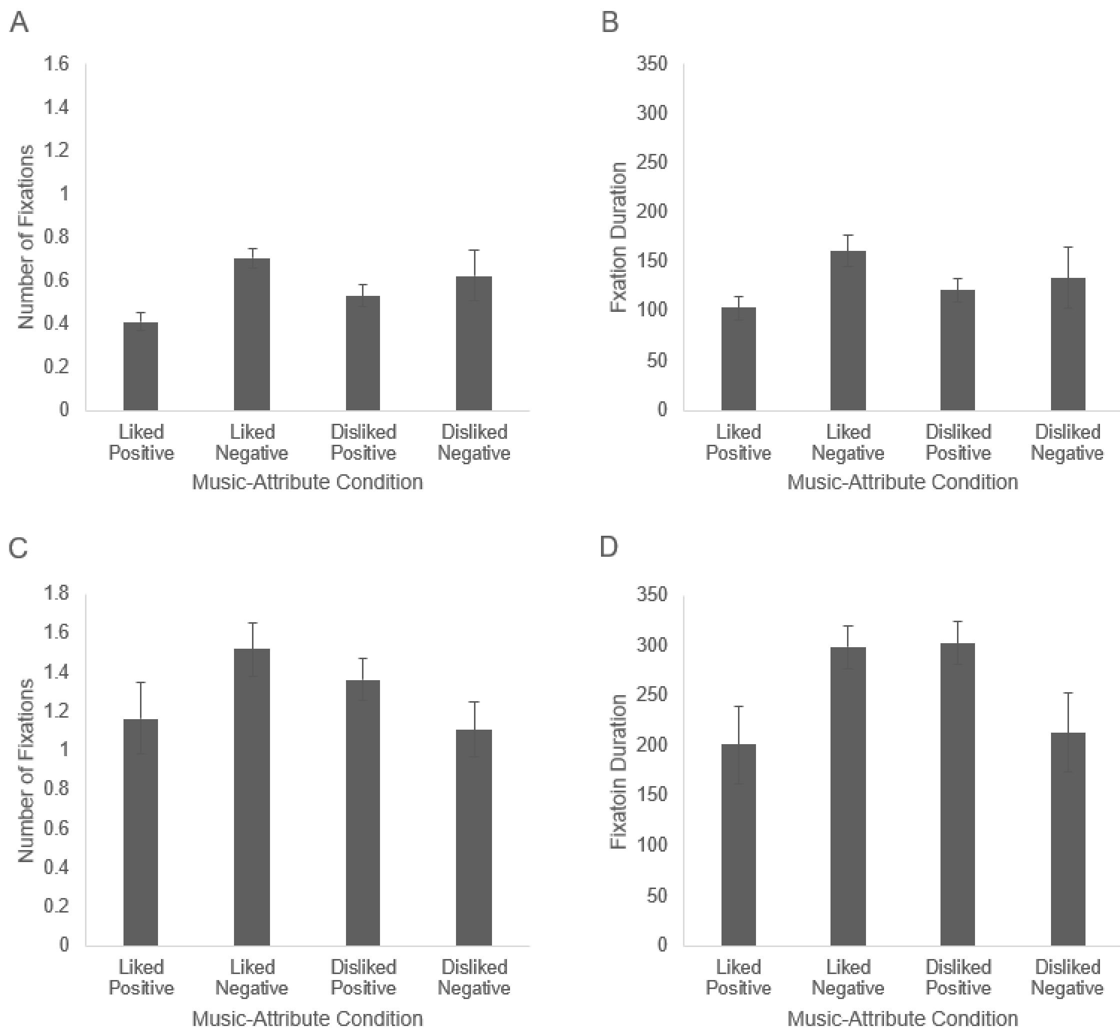

Figure 2.2 Looking measures for attributes in Study 1. Total number of fixations for positive and negative attributes paired with liked and disliked music during induction (Panel A) and choice (Panel C). Total fixation durations (in milliseconds) for positive and negative attributes paired with liked and disliked music during induction (Panel B) and choice (Panel D).

A parallel music valence (liked or disliked) × attribute valence (positive or negative) repeated-measures ANOVA was conducted on total fixation durations on the four attribute interest areas during induction (Figure 2.2, Panel B). There was no significant music valence × attribute valence interaction, $F(1,24) = 2.705$, $p = .113$, no main effect of attribute valence, $F(1,24) = 2.448$, $p = .131$, and no main effect of music valence, $F(1,24) = 0.063$, $p = .803$. Total looking time did not differ between conditions (Table 2.2).

Table 2.2 Descriptive statistics for eye tracking measures in the induction and paired choice phases of Study 2.

Phase	Total Number of Fixations		Total Fixation Duration (ms)	
Induction	Liked Music	Disliked Music	Liked Music	Disliked Music
Positive Attributes	0.410 (0.211)	0.530 (0.589)	103.636 (79.932)	122.153 (154.722)
Negative Attributes	0.704 (0.216)	0.622 (0.252)	161.210 (57.511)	134.085 (60.571)
Paired Choice				
Positive Attributes	1.255 (0.586)	1.426 (0.673)	200.415 (101.406)	327.620 (167.288)
Negative Attributes	1.605 (0.806)	1.148 (0.487)	319.895 (169.029)	221.177 (95.325)

Note. Means (standard deviations). ms = milliseconds

A third repeated-measures ANOVA with factors music valence (liked or disliked) × attribute valence (positive or negative) was conducted on first looks at the four attributes interest areas during induction. First look was defined as the location of the first fixation on the attribute interest areas, if any. There was a significant music valence × attribute valence interaction, $F(1,24) = 7.696$, $p = .011$, and a significant main effect of attribute valence, $F(1,24) = 32.282$, $p < .001$. The main effect of music valence was not significant, $F(1,24) = 0.005$, $p = .945$. For pictures presented with liked music, negative attributes (*M*

= 17.32, *SD* = 12.40) received more first looks than positive attributes (*M* = 7.88, *SD* = 4.99). For pictures presented with disliked music, negative attributes (*M* = 15.84, *SD* = 11.43) again received more first looks than positive attributes (*M* = 9.28, *SD* = 6.59).

Another repeated-measures ANOVA with factors music valence (liked or disliked) × attribute valence (positive or negative) was conducted on pupil diameters for fixations on the four attributes interest areas during induction. There were no significant effects, with pupil diameters similar for positive (*M* = 1082.50, *SD* = 257.16) and negative (*M* = 1078.51, *SD* = 268.66) attributes of pictures presented with liked music and positive (*M* = 1083.34, *SD* = 277.47) and negative (*M* = 1053.27, *SD* = 253.09) attributes of pictures presented with disliked music.

Parallel repeated-measures ANOVAs were conducted on eye movement measures during the paired choice test phase testing the hypothesis that during paired choice, total number of fixations, total fixation durations, and first looks would be greater for affectively congruent attributes. A music valence (liked or disliked) × attribute valence (positive or negative) repeated-measures ANOVA was conducted on total number of fixations on the four attribute interest areas during paired choice (Figure 2.2, Panel C). There was a significant music valence × attribute valence interaction, $F(1,17) = 17.798$, $p = .001$, and a significant main effect of music valence, $F(1,17) = 7.009$, $p = .017$. The main effect of attribute valence was not significant, $F(1,17) = 1.760$, $p = .202$. For pictures presented with liked music, negative attributes (*M* = 1.605) were looked at significantly more often than positive attributes (*M* = 1.255). For pictures presented with disliked music, positive attributes (*M* = 1.426) were looked at significantly more often than negative attributes (*M* = 1.148) (Table 2.2).

A parallel music valence (liked or disliked) × attribute valence (positive or negative) repeated-measures ANOVA was conducted on total fixation durations on the four attribute interest areas during paired choice (Figure 2.2, Panel D). There was a significant music valence × attribute valence interaction, $F(1,17) = 23.452$, $p < .001$. There was also a significant main effect of attribute valence, $F(1,17) = 5.200$, $p = .036$. There was no main effect of music valence, $F(1,17) = 0.348$, $p = .563$. For pictures presented with liked music, negative attributes ($M = 319.895$ ms) were looked at significantly longer than positive attributes ($M = 200.415$ ms). For pictures presented with disliked music, positive attributes ($M = 327.620$ ms) were looked at significantly longer than negative attributes ($M = 221.177$ ms) (Table 2.2).

A third repeated-measures ANOVA with factors music valence (liked or disliked) × attribute valence (positive or negative) was conducted on first looks at the four attributes interest areas during paired choice. First look was defined as the location of the first fixation on the attribute interest areas, if any. There was a significant music valence × attribute valence interaction, $F(1,17) = 24.14$, $p < .001$. There was no significant main effect of picture valence, $F(1,17) = 1.621$, $p = .221$, or of attribute valence, $F(1,17) = 2.137$, $p = .163$. For pictures presented with liked music, negative attributes ($M = 37.18$, $SD = 16.33$) received more first looks than positive attributes ($M = 26.53$, $SD = 11.74$). For pictures presented with disliked music, positive attributes ($M = 32.94$, $SD = 14.82$) received more first looks than positive attributes ($M = 27.18$, $SD = 11.25$).

Another repeated-measures ANOVA with factors music valence (liked or disliked) × attribute valence (positive or negative) was conducted on pupil diameters for fixations on the four attributes interest areas during paired choice. There were no significant effects, with pupil diameters similar for positive (*M* = 1109.34, *SD* = 261.49) and negative (*M* = 1120.63, *SD* = 264.32) attributes of pictures presented with liked music and positive (*M* = 1098.31, *SD* = 261.52) and negative (*M* = 1109.88, *SD* = 255.76) attributes of pictures presented with disliked music.

Additional repeated-measures ANOVAs were conducted on eye movement measures for the product pictures, testing the hypothesis that during paired choice, total number of fixations and total fixation duration would be greater for products paired with liked music than products paired with disliked music. A music valence (liked or disliked) × product category repeated-measures ANOVA was conducted on total number of fixations on the product pictures during paired choice. There was no main effect of music valence, $F(1, 24) = 0.645$, $p = .430$. There was a main effect of product category, $F(4,96) = 10.238$, $p < .001$, and no music valence × product category interaction, $F(4,96) = 0.337$, $p = .852$ (although these effects are not of primary interest). There was no difference in the number of looks for products paired with liked (*M* = 2.063, *SD* = 0.576) versus disliked (*M* = 2.026, *SD* = 0.628) music.

A second ANOVA with the same factors was conducted on total fixation duration on the product pictures during paired choice. There was no main effect of music valence, $F(1, 24) = 2.547$, $p = .096$. There was a main effect of product category, $F(4,96) = 5.793$, $p < .001$, and no music valence × product category interaction, $F(4,96) = 0.379$, $p = .823$. There was no significant difference in total looking time for products paired with liked (*M* = 480.656 ms, *SD* = 132.091 ms) versus disliked (*M* = 459.623 ms, *SD* = 142.298 ms) music.

Regression analyses on total number of fixations and total fixation duration tested whether these effects were moderated by contingency awareness, testing the hypothesis that greater contingency awareness would predict greater looking at affectively congruent features during conditioning. The dependent variables were calculated by subtracting looking at affectively incongruent attributes from looking at affectively congruent attributes separately by phase (induction or choice) and by measure (total number of fixations or total fixation duration). Contingency awareness was only a significant predictor of the average difference in total number of fixations for congruent versus incongruent attributes during paired choice, r = -.436, R^2 = .190, $F(1,19)$ = 4.452, p = .048, standardized β = -0.436 (in the opposite direction as predicted, Figure 2.3). One data point was observed to be an outlier (studentized residual = -2.99); conducting the same regression analysis without this point resulted in a slightly weaker but similar relation, R^2 = .187, $F(1,18)$ = 4.136, p = .057, standardized β = -0.432. Contingency awareness was not significantly related to total fixation duration during paired choice (r = -.200) nor to total number of fixations (r = -.108) or total fixation duration (r = -.041) during induction.

Supplementary analyses were conducted by participant group, with group categorization based on the tests of effects on paired choice at the individual level (significant effect in the predicted direction, significant effect in the reverse direction, no significant effect). These results are presented in Appendix A.

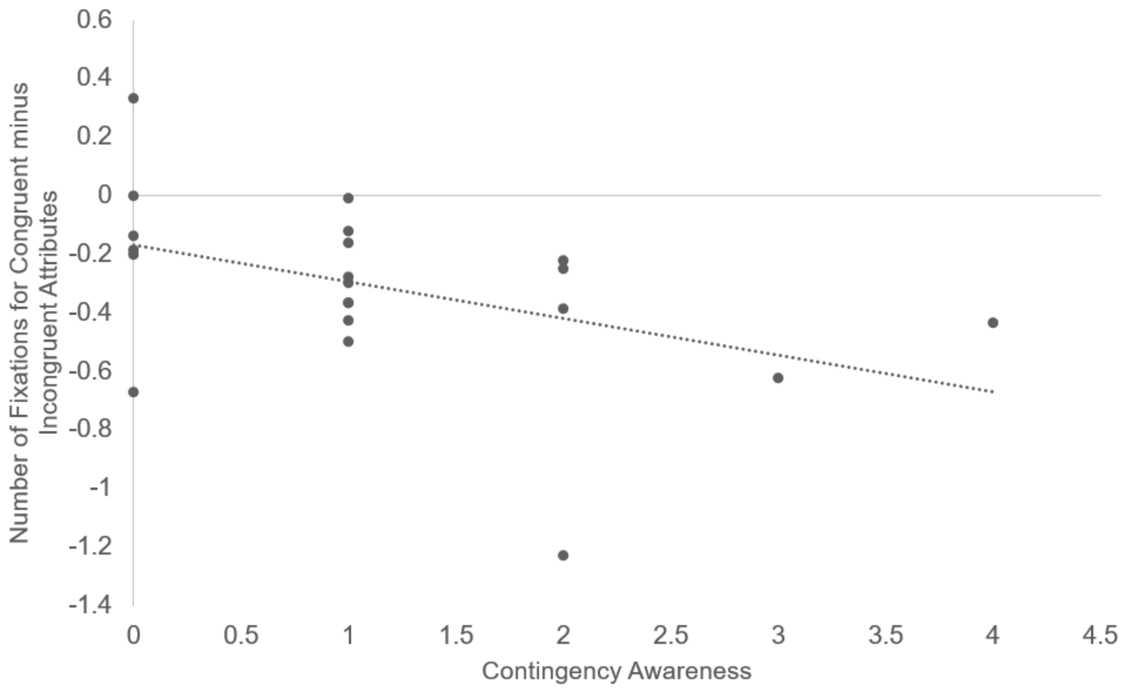

Figure 2.3 Contingency awareness and number of fixations in Study 1. Contingency awareness was scored as number correct out of 20. Total number of fixations shown for congruent minus incongruent attributes during paired choice.

2.3 DISCUSSION

Study 1 tested how visual attention to positive and negative features of stimuli during conditioning and testing was altered by the valence of the unconditioned stimulus, and how contingency awareness was related to these effects. While the effects of conditioning on behavioral and eye tracking measures were highly idiosyncratic and revealed strong individual differences, there are some consistent patterns of effects apparent in the results.

A sizeable proportion of participants showed choice preferences in the expected direction, choosing products paired with liked music significantly more often. A few participants showed a significant effect in the reverse direction, and another group showed no significant effects. This pattern of behavioral results and the lack of significant effect at the group level is consistent with previous research using the same conditioned product stimuli (Weber, 2018) and completely different stimuli (Weber et al., 2020). However, the power to detect significant effects was reduced by the reduction in sample size from the original goal, resulting in some ambiguity in interpreting these effects.

Looking behavior toward product attributes tended to show greater visual attention was given to features inconsistent with the valence of the paired music. Overall, participants looked more often and longer at negative (one- and two-star) attributes of products paired with liked music and positive (four- and five-star) attributes of products paired with disliked music. These attributes were also more likely to be looked at first, a potential indicator of increased saliency. These effects were strong during paired choice and somewhat reduced during induction.

This finding is actually the opposite of what was predicted, that total number of fixations, total fixation durations, and first looks would be greater for affectively congruent attributes. These hypotheses were formulated on the basis of the conceptual categorization account, which posits that conditioned stimuli are recategorized into a new valence category due to increased salience of affectively congruent features (Davey, 1994; De Houwer et al., 2001; Field & Davey, 1999). In Study 1, preferences for the conditioned stimuli did change for the majority of participants, but these conditioning effects were actually related to increased and earlier attention to incongruent attributes. Pairing the previously neutral products with the unconditioned music stimuli may have highlighted the oppositely charged attributes. This pattern of results therefore does not support the conceptual categorization account. Previously, the predictions of this model have received little attention in the literature, and the current results suggest that the more standard associative explanations are likely a better fit.

Effects on looking differentially toward attributes were stronger in paired choice than during induction. The task during induction may have reduced the likelihood of finding effects on attention to attributes during this phase. During induction, participants were required to answer a yes/no question about the picture of the product after each picture presentation. This task may have directed participants to attend only to the product pictures in expectation of these questions, leading to very little looking toward the attributes, as shown by the low means across all conditions. Future research should consider looking toward attributes using a task that focuses on the attributes or the unconditioned stimuli or using no task as a further test of the conceptual categorization account.

A significant relation was found between the effects of conditioning on attention to attributes during paired choice and participants' degree of contingency awareness. Higher levels of awareness were related to increased attention to affectively incongruent features. This finding is consistent with the role of an explicit mechanism in evaluative conditioning (Baeyens et al., 2005; Gast, 2018; Gawronski & Bodenhausen, 2006). However, contingency awareness was not related to choice behavior, although it was related to liking ratings in the subset of participants who showed a significant conditioning effect. The evidence for such explicit accounts is therefore limited in the current study. Many previous studies have found contingency awareness to be a significant predictor of conditioning (Blask et al., 2012; Field & Moore, 2005; Gast et al., 2012; Stahl et al., 2009).

Some of the nonsignificant findings in this study may be due to insufficient power. Although a power analysis was used to determine the appropriate sample size, the effect size used in the power analysis was based on experiments testing for evaluative conditioning on affective state responses rather than choice preferences. Choice preferences may be more likely to show individual differences resulting in a smaller overall effect at the group level (see Weber et al., 2020 versus Weber, 2018). Additionally, the sample size of 25 was not sufficiently powered to detect moderately sized correlations with contingency awareness. The original goal of 50 participants was unavoidably reduced due to the shutdown of experimental facilities as a result of COVID-19. Due to these circumstances, only 25 participants were recruited for this study, and Study 2 became a behavioral rather than eye tracking experiment as eye tracking was no longer feasible. Study 2 reexamines the role of contingency awareness with a larger sample. Study 2 also extends research into conditioning of ambivalent stimuli to those possessing more analogue positive and negative features and measures effects using liking ratings rather than choice behavior as the primary dependent variable.

CHAPTER 3

STUDY 2 – EVALUATIVE CONDITIONING WITH AMBIVALENT PICTURES

Study 2 was a behavioral experiment testing how responses to ambivalent picture stimuli are altered by evaluative conditioning. This study replicates some components of Study 1 using the same unconditioned music stimuli and different conditioned stimuli. In Study 2, the conditioned stimuli were ambivalent picture pairs created by combining positive and negative IAPS pictures. In a recent evaluative conditioning study, the same kind of positive-negative IAPS picture pairs were used as unconditioned stimuli (Glaser, Woud, Iskander, Schmalenstroth, & Vo, 2018). In Study 1, the star attribute ratings were symbolic representations of the valence characteristics of a stimulus, whereas in Study 2, the picture pairs have positive and negative components as a result of analog valence characteristics involving perceptual or semantic properties of the images (Paivio, 1991). To my knowledge, no study has tested for evaluative conditioning on conditioned stimuli designed to be ambivalent.

Another difference from Study 1 was that Study 2 measured conditioning effects using liking and disliking ratings separately rather than choice as the fundamental dependent variable. It was hypothesized that participants' liking ratings would be significantly higher and disliking ratings would be significantly lower for ambivalent stimuli paired with liked music than for ambivalent stimuli paired with disliked music. As in Study 1, Study 2 also tested whether contingency awareness moderated these effects using follow-up regression analyses.

3.1 METHOD

3.1.1 NORMING STUDIES

In a set of two norming studies, separate groups of participants made liking, disliking, and ambivalence ratings for positive, negative, and neutral IAPS pictures, as well as positive-negative and neutral-neutral picture pairs. The results from this pilot study were used to select picture pairs with consistently strongly liked and disliked halves to be used in the main experiment. These ambivalent stimuli for the main experiment were created from combinations of positive and negative IAPS pictures so that the positive half was strongly liked, the negative half was strongly disliked, and the degree of liking/disliking was well-matched between the two halves (an example is shown in Figure 3.1). In addition, four neutral-neutral stimuli pairs were created from components with neutral ratings (close to 5 on 9-point liking and disliking scales) to be used as control stimuli in the main experiment.

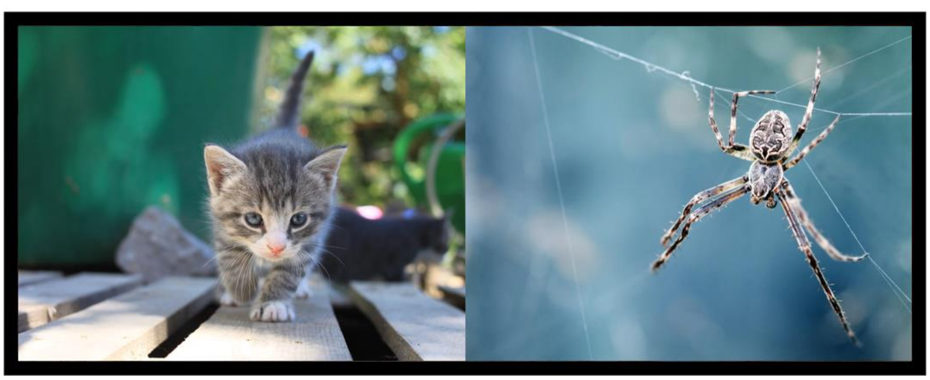

Figure 3.1 Example of an ambivalent picture pair. Example is similar to conditioned stimuli used in Study 2. Actual stimulus paired were created by combining a positive and a negative picture from the International Affective Picture System.

Liking and disliking ratings for the individual halves of the picture pairs were needed for two reasons. First, IAPs pictures are categorized as positive or negative based on ratings of emotional valence. In Study 2, we conditioned the liking for the stimuli and tested how liked and disliked music altered visual attention to different parts of the picture-pairs. It was therefore useful to create the picture pairs on the basis of liking ratings from our own study population, rather than on valence ratings from a different population. Second, one possibility is that liking and disliking rates for picture pairs are produced by averaging together the two halves, in which case we might expect liking and disliking ratings for ambivalent positive-negative picture pairs to be approximately equal to ratings for neutral-neutral picture pairs. Collecting ratings for the individual components allowed us to ensure ambivalent picture pairs had highly liked and disliked features, even though their holistic ratings may have been close to neutral.

3.1.2 PARTICIPANTS

We recruited 81 University of South Carolina undergraduates for this experiment using SONA. As described in Study 1, this sample size provided a power greater than .99 to detect the effect of conditioning (0.79) and a power greater than .80 to detect a moderately strong relation between evaluative conditioning effects and contingency awareness. Of the 81 participants, we excluded 9 who did not fully complete the study, 9 who were unable to hear the music stimuli due to an Internet browser compatibility issue, and 3 who failed a manipulation check for the music. The final sample (n = 60) comprised 49 women and 11 men, ages 18 – 56 (M = 20.48, SD = 4.93).

3.1.3 MATERIALS

The unconditioned stimuli were a subset of the liked and dislike music clips described in Study 1. There were eight liked music clips and eight disliked music clips. The conditioned stimuli were twelve ambivalent positive-negative picture pairs and four neutral-neutral picture pairs. Ambivalent stimuli were created by combining highly liked positive IAPS pictures (M = 8.22) with highly disliked negative IAPS pictures (M = 1.87). Neutral-neutral picture stimuli were created by combining two neutral IAPS pictures with neutral liking and disliking ratings (M = 4.26). Including neutral stimuli allowed for comparisons of conditioning effects on these neutral stimuli, which are typical in evaluative conditioning (Hofmann et al., 2010), to effects on ambivalent stimuli, which have not been used as conditioned stimuli in any previous study we are aware of. Each stimulus pair consisted of two pictures presented side by side surrounded by an outer border (as in Glaser et al., 2018). Six liked and six disliked music clips were paired with the ambivalent picture pairs, and two liked and two disliked music clips were paired with the neutral-neutral picture pairs.

3.1.4 PROCEDURE

Unlike Study 1, Study 2 was a web-based study administered using the Qualtrics testing platform. The study proceeded in four phases: induction, test, music rating, and contingency awareness measure (Table 3.1). During induction, participants viewed the picture stimuli paired simultaneously with either liked or disliked music. Each picture-music pairing was presented 8 times, for a total of 128 trials presented in a random order for each participant. After each presentation, participants answered a simple yes or no question about a perceptual feature of the picture pair. Features include red, yellow, green, blue, white, black, curved lines, and straight lines. For example, participants were asked, “Did you see any RED in the picture?”

In the test phase, participants made liking, disliking, and ambivalence ratings for each picture pair blocked by rating scale. There were 48 trials (16 exemplars × 3 ratings). Participants first rated how much they liked each picture pair on a Likert-type scale from 1 (*Not at All)* to 9 (*Very Much*). Second, they rated how much they disliked each picture pair from 1 (*Not at All*) to 9 (*Very Much*). Third, they rated how ambivalent each picture pair was from 1 (*Not at All*) to 9 (*Very Much*).

Table 3.1 Materials and procedures used in Study 2.

Induction	Test	Contingency Awareness Measure
Trials: 128 (2 music conditions × 8 exemplars × 8 repetitions) **Task:** Answer yes/no perceptual questions	**Trials:** 48 (2 music conditions × 8 exemplars × 3 ratings) **Task:** Rate liking, disliking, and ambivalence Followed by **music ratings** (liking)	**Trials:** 16 (2 music conditions × 8 exemplars) **Task:** Choose the picture that was presented with the music clip

Note. The 8 exemplars included 6 ambivalent and 2 neutral picture pairs.

In the music rating phase, participants rated liking for the music alone, serving as a manipulation check to verify the liking ratings of the unconditioned stimuli. Participants rated how much they liked each music clip on a Likert-type scale from 1 (*Not at All)* to 9 (*Very Much*).

In the contingency awareness measure, participants heard each music clip while seeing all 16 picture pairs on the screen and were asked to select the picture that was shown when they heard that music clip earlier.

3.2 RESULTS

3.2.1 MANIPULATION CHECKS

A paired-samples *t*-test was conducted on the liking ratings for the liked and disliked music to test the hypothesis that liked music would be rated as significantly more liked than disliked music. As expected, liked music (M = 7.71, SD = 0.93) was significantly more liked than disliked music (M = 1.50, SD = 0.88), $t(59) = 36.824$, $p < .001$ (Figure 3.2, Panel A). The difference in liking between liked and disliked music varied between 2.87 and 8.00 across participants (Figure 3.2, Panel B).

A repeated-measures ANOVA with factors music condition (liked or disliked) and stimulus type (ambivalent or neutral) was conducted on ambivalence ratings of the picture pairs to test the hypothesis that ambivalent picture pairs would be rated as significantly more ambivalent than neutral-neutral picture pairs. As expected, there was a significant main effect of stimulus type, $F(1,59) = 95.117$, $p < .001$. Ambivalent picture pairs (6.15) were rated as significantly more ambivalent than neutral-neutral picture pairs (3.13), $p < .001$.The stimulus type × music condition interaction effect was not significant, $F(1,59) = 0.127$, $p = .723$, nor was the main effect of music condition, $F(1,59) = 0.349$, $p = .557$.

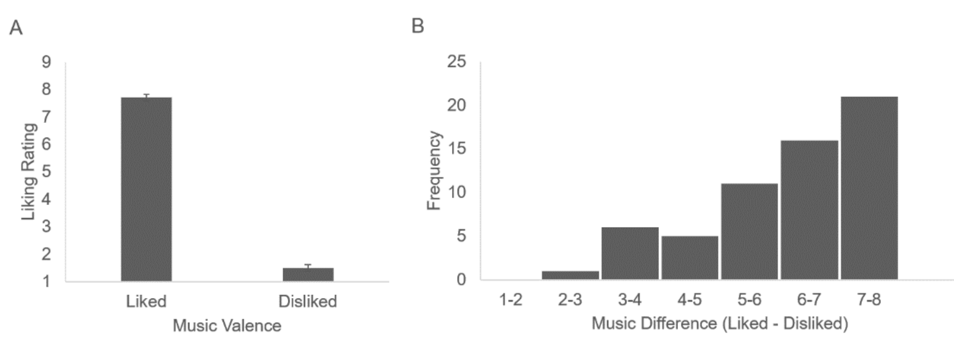

Figure 3.2 Liking ratings for music in Study 2. Panel A: mean liking for liked and disliked music with standard error bars. Panel B: variation in music liking between conditions across participants (calculated by subtracting mean liking for disliked music from mean liking for liked music).

3.2.2 EVALUATIVE CONDITIONING EFFECTS

Two repeated-measures ANOVAs were conducted to test the effects of liking and disliking ratings of the picture pairs.

First, a repeated-measures ANOVA with factors music condition (liked or disliked) and stimulus type (ambivalent or neutral) was conducted on liking ratings to test the hypothesis that stimuli paired with liked music would be rated as more liked than stimuli paired with disliked music. There was a significant main effect of music condition, $F(1,59) = 7.981$, $p = .006$ (Figure 3.3, Panel A). The music condition × stimulus type interaction effect was not significant, $F(1,59) = 0.010$, $p = .922$, nor was the main effect of stimulus type, $F(1,59) = 1.213$, $p = .275$. Ambivalent picture pairs were significantly more liked after being presented with liked music ($M = 4.75$, $SD = 1.35$) than with disliked music ($M = 4.30$, $SD = 1.32$), p = 005. The pairwise comparison between neutral-neutral picture pairs presented with liked music ($M = 4.51$, $SD = 1.75$) and disliked music ($M = 4.09$, $SD = 1.93$) was not significant, $p = .103$, but in the correct direction.

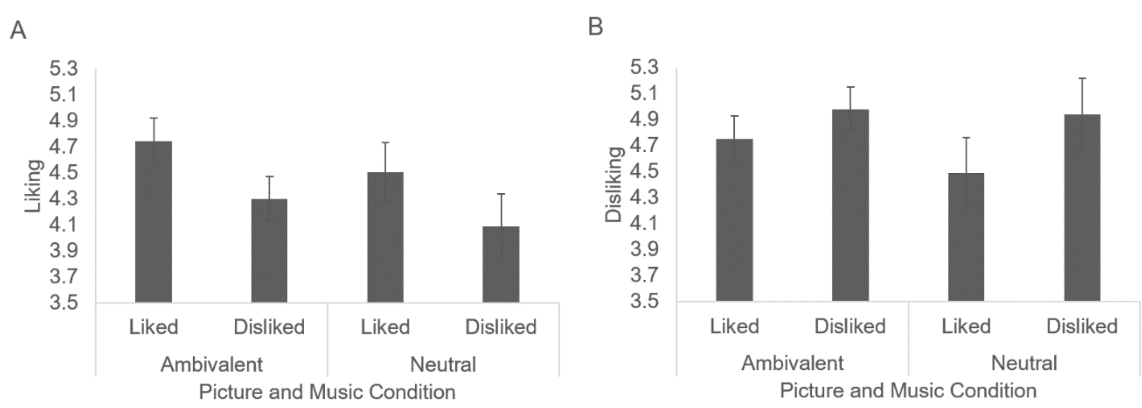

Figure 3.3 Liking and disliking ratings for pictures in Study 2. Liking (Panel A) and disliking (Panel B) ratings for ambivalent and neutral picture paired previously presented with liked or disliked music, with standard error bars.

Second, a repeated-measures ANOVA with factors music condition (liked or disliked) and stimulus type (ambivalent or neutral) was conducted on disliking ratings to test the hypothesis that stimuli paired with liked music would be rated as less disliked than stimuli paired with disliked music. The main effect of music condition was marginally significant, $F(1,59) = 3.312$, $p = .074$ (Figure 3.3, Panel B). The music condition × stimulus type interaction effect was not significant, $F(1,59) = 0.403$, $p = .528$, nor was the main effect of stimulus type, $F(1,59) = 0.507$, $p = .479$. The pairwise comparisons between ambivalent picture pairs presented with liked music ($M = 4.75$, $SD = 1.37$) and disliked music ($M = 4.98$, $SD = 1.33$) was not significant, $p = .184$. The pairwise comparison between neutral-neutral pictures presented with liked music ($M = 4.49$, $SD = 2.09$) and disliked music ($M = 4.94$, $SD = 2.13$) was also not significant, $p = .164$.

By-item paired-samples *t*-tests were conducted to further determine whether liking and disliking ratings differed for pictures paired with liked or disliked music. Pictures paired with liked music ($M = 4.69$, $SD = 0.54$) were significantly more liked than pictures paired with disliked music ($M = 4.24$, $SD = 0.56$), $t(15) = 3.196$, $p = .006$. Pictures paired with disliked music ($M = 4.98$, $SD = 0.52$) were significantly more disliked than pictures paired with liked music ($M = 4.68$, $SD = 0.64$), $t(15) = 2.554$, $p = .022$.

A linear regression analysis was used to test whether music difference scores (calculated by subtracting liking ratings for disliked music from ratings for liked music) predicted evaluative conditioning effects on liking in the test phase (calculated by taking difference scores for liking ratings of stimuli paired with liked versus disliked music). Music difference scores were not correlated with and did not predict evaluative conditioning effects, $r = .072$, $F(1,58) = 0.306$, $R^2 = .005$, standardized $\beta = 0.072$, $p = .582$.

3.2.3 CONTINGENCY AWARENESS

Three contingency awareness scores were calculated for each participant. An "identity" score was calculated as the number of trials out of 16 on which a participant correctly identified the exact picture which had been presented with the music clip. A "matching" score was calculated as the number of trials out of 16 on which a participant chose a picture which had been presented with a music clip of the same condition (liked or disliked) as the music clip presented. A "semantic" score was calculated as the number of trials out of 16 on which a participant chose a picture which matched the semantic category of the correct picture (animal, human, or scene). The mean identity score was 3.88 (*SD* = 2.44, range 0 – 11), the mean matching score was 11.42 (*SD* = 2.67, range 5 – 16), and the mean semantic score was 6.97 (*SD* = 2.30, range 2-12).

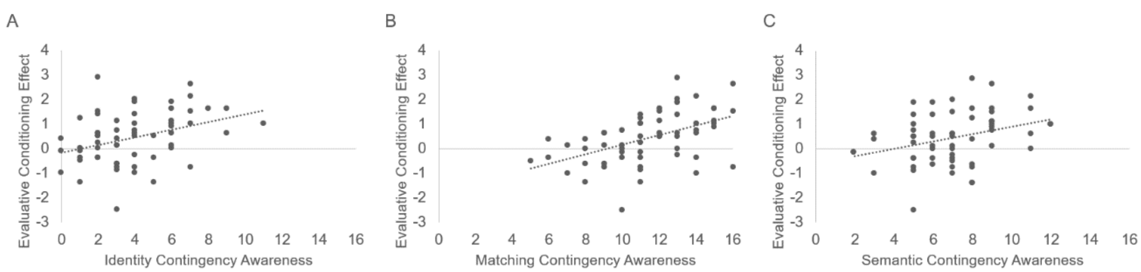

Figure 3.4 Contingency awareness and evaluative conditioning effects in Study 2. Relations between identity contingency awareness (calculated as number correct out of 16, Panel A), matching contingency awareness (calculated as number with matching valence out of 16, Panel B), and semantic contingency awareness (calculated as number in correct semantic category, Panel C) and evaluative conditioning effects (calculated by subtracting liking scores for pictures paired with disliked music minus liking for pictures paired with liked music).

Linear regression analyses were used to test whether identity, matching, or semantic contingency awareness scores predicted evaluative conditioning effects on liking in the test phase (calculated by taking difference score for liking ratings of stimuli paired with liked versus disliked music). These analyses tested the hypothesis that contingency awareness would predict evaluative conditioning effects, with greater awareness predicting larger effects. Identity scores significantly predicted liking difference scores, $F(1,58) = 8.832$, $R^2 = .132$, standardized $\beta = 0.364$, $p = .004$. Matching scores also significantly predicted liking difference scores, $F(1,58) = 18.607$, $R^2 = .243$, standardized $\beta = 0.493$, $p < .001$. Finally, semantic scores also significantly predicted liking difference scores, $F(1,58) = 7.059$, $R^2 = .108$, standardized $\beta = 0.329$, $p = .010$.

3.3 DISCUSSION

Study 2 serves to further elucidate the mechanisms underlying evaluative conditioning by examining how evaluative conditioning generalizes to ambivalent stimuli with more analogue features. It also tested how contingency awareness was related to these effects.

Using a larger sample size, analyses were better powered to detect effects. There was a significant effect of conditioning at the group level, with pictures previously presented with liked music rated as more highly liked. These results serve to reinforce the power of the unconditioned music stimuli in producing conditioning effects. Unlike Study 1, there was a significant effect of music condition at the group level. One explanation for this difference may be the change in the conditioned stimulus format or in the primary dependent variable. Whereas Study 1 tested effects primarily on forced choices, Study 2 tested effects on liking ratings of the stimuli. Most studies of evaluative conditioning use Likert-type ratings rather than forced-choice procedures to assess changes in preference (Jones, Olson & Fazio, 2010). Study 2 was also more strongly powered than Study 1, which may have contributed to this difference between experiments. The effect of conditioning on disliking ratings did not reach significance in the by-participants analysis, although it was significant in the by-items analysis. Disliking ratings always occurred second following the liking ratings, which may have resulted in less extreme evaluations.

The effects of conditioning on ambivalent picture pairs were similar to those observed for more traditional neutral picture pairs. To my knowledge, this is the first study demonstrating evaluative conditioning on stimuli designed to be ambivalent. This finding serves to expand the external validity of evaluative conditioning. In real-world situations, many stimuli contain both positive and negative features. For example, advertisements may attempt to shift a negative attitude toward a product or public figure to be more positive by pairing them with a liked stimulus. Scenes in television and movies with multiple characters may be similarly ambivalent. Some IAPS pictures treated as neutral in past experiments are actually both positive and negative, an ambivalence that results in ratings close to neutral when tested on a single scale (Schneider et al., 2016). The current study shows that evaluative conditioning can successfully alter responses to stimuli possessing both positive and negative components. In addition, the observation of these conditioning effects in an online study rather than a traditional laboratory setting serves to demonstrate the robustness of these effects in different contexts.

These effects were also predicted by two measures of contingency awareness. This finding is in line with many previous studies showing that conditioning effects are dependent on awareness (Blask et al., 2012; Field & Moore, 2005; Gast et al., 2012; Stahl et al., 2009). With the larger sample size in Study 2 compared to Study 1, analyses were sufficiently powered to detect these effects. These results are consistent with the role of a conscious, explicit mechanism in evaluative conditioning. The implicit process models would argue that the magnitude of conditioning effects should not depend on participants' knowledge of the stimulus pairings (Baeyens et al., 1992; Field & Davey, 1999; Jones et al., 2009; Martin & Levey, 1994). This finding is instead consistent with predictions of the propositional model (Baeyens et al. 2005), the Declarative Memory Model (Gast, 2018), and the APE model (Gawronski & Bodenhausen, 2006).

The conclusions regarding the role of contingency awareness are somewhat limited as awareness was only measured and not manipulated. Previous literature has demonstrated that measurements of contingency awareness may be confounded with memory (Gawronski & Walther, 2012). Future research should manipulate contingency awareness or measure it during learning to avoid this potential confound.

CHAPTER 4

GENERAL DISCUSSION

Study 1 tested a specific instantiation of the predictions of the conceptual categorization model, a non-associative account of evaluative conditioning. Both Studies 1 and 2 tested for the role of contingency awareness in evaluative conditioning. Furthermore, these studies generalized the evaluative conditioning paradigm to stimuli designed to be ambivalent rather stimuli selected to be neutral. Together, these studies use behavioral and eye tracking approaches to provide insights into the mechanisms underlying evaluative conditioning, whether they be non-associative or associative and implicit or explicit.

4.1 THE CONCEPTUAL CATEGORIZATION MODEL

The conceptual categorization model was tested by investigating how attention to visual attention to positive and negative features was altered by conditioning in Study 1. Unlike most models of evaluative conditioning, the conceptual categorization account proposes that conditioning effects are non-associative and are driven by a recategorization of the conditioned stimulus as a result of increased saliency of features affectively congruent with the unconditioned stimulus (Field & Davey, 1999). The observed findings are counter to the predictions of this model. Instead of increasing visual attention to affectively congruent features, the overall pattern of results showed increased attention to affectively incongruent features. This finding builds on prior research that tested for changes in looking during testing and extends to changes in looking during conditioning. There was no evidence for increased visual saliency of congruent features as measured by looking time in either induction or paired choice. In fact, evidence for the opposite effects were found, with greater and earlier looking at incongruent features. This result is difficult to explain under the assumptions of the conceptual categorization account, suggesting the effects of evaluative conditioning may be underlain by more traditional associative processes.

4.2 CONTINGENCY AWARENESS

The results of the current studies offer somewhat mixed evidence for the role of contingency awareness in evaluative conditioning. Although Study 1 did not find significant effects of awareness on conditioning in the full sample, the ability to detect this effect was limited by the power of the study and thus the null effect is difficult to interpret. Awareness did predict the change in product liking ratings in a subset of participants. Study 2, which was more strongly powered for correlational analyses, found that three measures of contingency awareness significantly predicted conditioning effects on liking.

These results suggest that a conscious mechanism may be involved in evaluative conditioning and are consistent with many previous studies showing that conditioning effects are stronger with greater awareness or even present only under aware conditions (Blask et al., 2012; Field & Moore, 2005; Gast, et al., 2012; Stahl et al., 2009). Participants who were able to display greater knowledge of the stimulus pairings were more likely to show significant behavioral effects in the predicted direction. The explicit and dual-process models of evaluative conditioning would argue that participants' evaluations of the ambivalent stimuli during the test phases were created using this conscious knowledge

(Baeyens et al., 2005; Gast, 2018; Gawronski & Bodenhausen, 2006). However, it is possible that although participants had knowledge about the stimulus pairings, this knowledge was not actually used in producing stimulus evaluations. The implicit process models would argue that the effects of conditioning were driven by the strength of associative links between the unconditioned and conditioned stimuli, and contingency awareness is a by-product of conditioning rather than a driving force.

In these experiments, contingency awareness was measured rather than manipulated. Because awareness was tested at the end of the experiment, the measure was confounded with the effects of memory (Gawronski & Walther, 2012). For example, some participants may have had more explicit knowledge of the stimulus pairings during the earlier paired choice phase, but some information decayed by the time of the contingency awareness measure. Furthermore, it is possible that simply measuring awareness heightened participants' conscious knowledge of the stimuli by requiring them to effortfully recall the pairings. Manipulating awareness with a task that increases cognitive load during conditioning could yield more insights to the role of explicit knowledge in evaluative conditioning as a more powerful test of the explicit and dual-process model predictions.

4.3 AMBIVALENT STIMULI IN EVALUTIVE CONDITIONING

These findings also generalize evaluative conditioning to two types of stimuli designed to be ambivalent. In Study 1, ambivalent stimuli were created by presenting pictures of consumer products with positive and negative star ratings on several attributes. Pairing these ambivalent products with liked and disliked music produced significant conditioning effects for the majority of participants. In Study 2, ambivalent stimuli were paired combinations of liked and disliked IAPS photographs. Pairing these pictures with music resulted in conditioning at the group level. Together, this pair of studies demonstrates that evaluative conditioning can successfully extend from altering responses to neutral stimuli to those to ambivalent stimuli.

4.4 LIMITATIONS AND FUTURE DIRECTIONS

There are several limitations and opportunities for future research into the mechanisms of evaluative conditioning. First, Study 1 was unfortunately limited by its sample size due to the closing of experimental facilities as a result of COVID-19, reducing the power to detect significant effects both in conditioning and in the effect of contingency awareness on conditioning. Study 2 helped to address these limitations by increasing the sample size, revealing significant effects at the group level and the predictive power of contingency awareness, although it was a behavioral rather than eye tracking experiment due to the effects of COVID-19.

Second, Study 1 represents only one instantiation of the predictions of the conceptual categorization account, and other tests of this model are possible. One important limitation in Study 1 was that participants' attention was directed to the product pictures and away from the positive and negative attributes during learning by requiring them to answer questions about perceptual features of the products. As the average total number of fixations and average total fixation durations show, this reduced visual attention to the attributes during learning compared to test. Future research should consider different tasks that will not direct attention away from the valenced attributes.

Another possible test of the conceptual categorization model could use eye tracking examine attention to the liked and disliked components of the ambivalent picture pairs used in Study 2. Although the behavioral approach used here extended evaluative conditioning effects to these stimuli, eye tracking would allow for better insight into the attentional mechanisms that may underlie these effects. Based on the results of Study 1, it is reasonable to expect that participants may attend more to the portions of the picture pairs that are incongruent with the music. Study 2 demonstrates the feasibility of using ambivalent picture stimuli in investigating these effects.

Third and finally, the role of contingency awareness should be further studied by both directly manipulating awareness and by measuring awareness during learning rather than post-learning. Awareness could be manipulated by varying the task during learning across or within participants. For example, manipulating cognitive load during conditioning by requiring participants to complete an n-back task would likely reduce their explicit knowledge of the stimulus pairings by focusing attention elsewhere. Measuring awareness during learning would help eliminate the memory confound when measuring awareness at a later point (Gawronski & Walther, 2012).

4.5 CONCLUSION

This set of experiments examined visual attention to stimulus features in evaluative conditioning and how evaluative conditioning extends to ambivalent stimuli. Counter to the predictions of the conceptual categorization model, pairing of ambivalent conditioned stimuli with liked or disliked unconditioned stimuli did not result in greater attention to affectively congruent features. Instead, the overall pattern of results showed enhanced attention to affectively incongruent features. Study 2 but not Study 1 showed that contingency awareness was a significant predictor of conditioning effects, indicating that conscious knowledge plays a role in evaluative conditioning. Together, the findings of these studies make a significant contribution to our understanding of the role of attention and awareness in evaluative conditioning. Based on the present findings, there is no evidence that the unconditioned stimuli lead to greater processing of affectively congruent features of the ambivalent stimuli. However, conditioning effects for these ambivalent stimuli were related to participants' conscious awareness of the relations between the stimuli.

REFERENCES

Baeyens, F., Eelen, P., Crombez, G., & Van den Bergh, O. (1992). Human evaluative conditioning: Acquisition trials, presentation schedule, evaluative style and contingency awareness. *Behaviour Research and Therapy*, *30*(2), 133-142. https://doi.org/10.1016/0005-7967(92)90136-5

Baeyens, F., Eelen, P., & Van den Bergh, O. (1990). Contingency unawareness in evaluative conditioning: A case for unaware affective-evaluative learning. *Cognition & Emotion, 4*(1), 3-18. https://doi.org/10.1080/02699939008406760

Baeyens, F., Eelen, P., Van den Bergh, O., & Crombez, G. (1989). The influence of CS-UCS perceptual similarity/dissimilarity on human evaluative learning and signal learning. *Learning and Motivation, 20*(3), 322-333. https://doi.org/10.1016/0023-9690(89)90011-8

Baeyens, F., Field, A. P., and De Houwer, J. (2005). Associative learning of likes and dislikes: Some current controversies and possible ways forward. *Cognition & Emotion*, *19*(2), 161-174. https://doi.org/10.1080/02699930441000265

Bar-Anan, Y., De Houwer, J., & Nosek, B. A. (2010). Evaluative conditioning and conscious knowledge of contingencies: A correlational investigation with large samples. *The Quarterly Journal of Experimental Psychology, 63*(12), 2313-2335. https://doi.org/10.1080/17470211003802442

Blask, K., Walther, E., Halbeisen, G., & Weil, R. (2012). At the crossroads: Attention, contingency awareness, and evaluative conditioning. *Learning and Motivation, 43*(3), 99-106. https://doi.org/10.1016/j.lmot.2012.03.004

Davey, G. C. (1994). Defining the important theoretical questions to ask about evaluative conditioning: A reply to Martin and Levy (1994). *Behaviour Research and Therapy*, *32*(3), 307-310. https://doi.org/10.1016/0005-7967(94)90126-0

De Houwer, J. (2006). Using the Implicit Association Test does not rule out an impact of conscious propositional knowledge on evaluative conditioning. *Learning and Motivation, 37*(2), 176-187. https://doi.org/10.1016/j.lmot.2005.12.002

De Houwer, J. (2007). A conceptual and theoretical analysis of evaluative conditioning. *The Spanish Journal of Psychology, 10*(2), 230-241. https://doi.org/10.1017/S1138741600006491

De Houwer, J., Thomas, S., & Baeyens, F. (2001). Association learning of likes and dislikes: A review of 25 years of research on human evaluative conditioning. *Psychological Bulletin, 127*(6), 853-869. https://doi.org/10.1037/0033-2909.127.5.853

Field, A. P., & Davey, G. C. (1999). Reevaluating evaluative conditioning: A nonassociative explanation of conditioning effects in the visual evaluative conditioning paradigm. *Journal of Experimental Psychology: Animal Behavior Processes*, *25*(2), 211-224. https://doi.org/10.1037/0097-7403.25.2.211

Field, A. P., & Moore, A. C. (2005). Dissociating the effects of attention and contingency awareness on evaluative conditioning effects in the visual paradigm. *Cognition & Emotion, 19*(2), 217-243. https://doi.org/10.1080/02699930441000292

Gast, A. (2018). A declarative memory model of evaluative conditioning. *Social Psychological Bulletin, 13,* e28590. https://doi.org/10.5964/spb.v13i3.28590

Gast, A., De Houwer, J., & De Schryver, M. (2012). Evaluative conditioning can be modulated by memory of the CS-US pairings at the time of testing. *Learning and Motivation, 43*(3), 116-126. https://doi.org/10.1016/j.lmot.2012.06.001

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin, 132*(5), 692-731. https://doi.org/10.1037/0033-2909.132.5.692

Gawronski, B., & Bodenhausen, G. V. (2011). The associative-propositional evaluation model: Theory, evidence, and open questions. In *Advances in experimental social psychology* (Vol. 44, pp. 59-127). Academic Press. https://doi.org/10.1016/B978-0-12-385522-0.0002-0

Gawronski, B., & Mitchell, D. G. (2014). Simultaneous conditioning of valence and arousal. *Cognition & Emotion, 28*(4), 577-595. https://doi.org/10.1080/02699931.2013.843506

Gawronski, B., & Walther, E. (2012). What do memory data tell us about the role of contingency awareness in evaluative conditioning? *Journal of Experimental Social Psychology*, *48*(3), 617-623. https://doi.org/10.1016/j.jesp.2012.01.002

Glaser, T., Woud, M. L., Iskander, M. L., Schmalenstroth, V., & Vo, T. M. (2018). Positive, negative, or all relative? Evaluative conditioning of ambivalence. *Acta Psychologica*, *185*, 155-165. https://doi.org/10.1016/j.actpsy.2018.02.006

Hofmann, W., De Houwer, J., Perugini, M., Baeyens, F., & Crombez, G. (2010). Evaluative conditioning in humans: A meta-analysis. *Psychological Bulletin*, *136*(3), 390-421. https://doi.org/10.1037/a0018916

Hütter, M., & Sweldens, S. (2013). Implicit misattribution of evaluative responses: Contingency-unaware evaluative conditioning requires simultaneous stimulus presentations. *Journal of Experimental Psychology: General*, *142*(3), 638-643. https://doi.org/10.1037/a0029989

Hütter, M., Sweldens, S., Stahl, C., Unkelbach, C., & Klauer, K. C. (2012). Dissociating contingency awareness and conditioned attitudes: Evidence of contingency-unaware evaluative conditioning. *Journal of Experimental Psychology: General*, *141*(3), 539-557. https://doi.org/10.1037/a0026477

Jones, C. R., Fazio, R. H., & Olson, M. A. (2009). Implicit misattribution as a mechanism underlying evaluative conditioning. *Journal of Personality and Social Psychology*, *96*(5), 933-948. https://doi.org/10.1037/a0014747

Jones, C. R., Olson, M. A., & Fazio, R. H. (2010). Evaluative conditioning: The “how” question. *Advances in Experimental Social Psychology, 43*, 205-255. https://doi.org/10.1016/S0065-2601(10)43005-1

Kardes, F. R., Posavac, S. S., & Cronley, M. L. (2004). Consumer inference: A review of processes, bases, and judgment contexts. *Journal of Consumer Psychology*, *14*(3), 230-256. https://doi.org/10.1207/s15327663jcp1403_6

Keating, G. D. (2014). Eye-tracking with text. In J. Jegerski & B. VanPatten (Eds.), *Research Methods in Second Language Psycholinguistics* (pp. 69-92). London: Taylor & Francis.

Martin, I., & Levey, A. (1994). The evaluative response: Primitive but necessary. *Behaviour Research and Therapy, 32*(3), 301-305. https://doi.org/10.1016/0005-7967(94)90125-2

Martin, I., & Levey, A. B. (1978). Evaluative conditioning. *Advances in Behaviour Research and Therapy, 1*(2), 57-101. https://doi.org/10.1016/0146-6402(78)90013-9

Mitchell, C. J., De Houwer, J., & Lovibond, P. F. (2009). The propositional nature of human associative learning. *Behavioral and Brain Sciences, 32*(2), 183-198. https://doi.org/10.1017/S0140525X09000855

Olson, M. A., & Fazio, R. H. (2006). Reducing automatically activated racial prejudice through implicit evaluative conditioning. *Personality and Social Psychology Bulletin, 32*(4), 421-433. https://doi.org/10.1177/0146167205284004

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, *45*(3), 255-287. http://doi.org/10.1037/h0084295

Perham, N., & Sykora, M. (2012). Disliked music can be better for performance with liked music. *Applied Cognitive Psychology, 26*(4), 550-555. https://doi.org/10.1002/acp.2826

Pleyers, G., Corneille, O., Luminet, O., & Yzerbyt, V. (2007). Aware and (dis) liking: Item-based analyses reveal that valence acquisition via evaluative conditioning emerges only when there is contingency awareness. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(1), 130-144. https://doi.org/10.1037/0278-7393.33.1.130

Priester, J. R., & Petty, R. E. (1996). The gradual threshold model of ambivalence: Relating the positive and negative bases of attitudes to subjective ambivalence. *Journal of Personality and Social Psychology, 71*(3), 431-449. http://doi.org/10.1037/0022-3514.71.3.431

Schneider, I. K., Veenstra, L., van Harreveld, F., Schwarz, N., & Koole, S. L. (2016). Let's not be indifferent about neutrality: Neutral ratings in the International Affective Picture System (IAPS) mask mixed affective responses. *Emotion, 16*(4), 426-430. https://doi.org/10.1037/emo0000164

Stahl, C., & Unkelbach, C. (2009). Evaluative learning with single versus multiple unconditioned stimuli: The role of contingency awareness. *Journal of Experimental Psychology: Animal Behavior Processes, 35*(2), 286-291. https://doi.org/10.1037/a0013255

Stahl, C., Unkelbach, C., & Corneille, O. (2009). On the respective contributions of awareness of unconditioned stimulus valence and unconditioned stimulus identity in attitude formation through evaluative conditioning. *Journal of Personality and Social Psychology, 97*(3), 404-420. https://doi.org/10.1037/a0016196

Stuart, E. W., Shimp, T. A., & Engle, R. W. (1987). Classical conditioning of consumer attitudes: Four experiments in an advertising context. *Journal of Consumer Research*, *14*(3), 334-349. https://doi.org/10.1086/209117

Walther, E. (2002). Guilty by mere association: Evaluative conditioning and the spreading attitude effect. *Journal of Personality and Social Psychology*, *82*(6), 919-934. https://doi.org/10.1037/0022-3514.82.6.919

Walther, E., & Nagengast, B. (2006). Evaluative conditioning and the awareness issue: Assessing contingency awareness with the four-picture recognition test. *Journal of Experimental Psychology: Animal Behavior Processes*, *32*(4), 454-459. https://doi.org/10.1037/0097-7403.32.4.454

Weber, C. E. (2018). *Evaluative conditioning of product preferences: An eye-tracking perspective.* Unpublished master's thesis, University of South Carolina, Columbia, South Carolina.

Weber, C. E., Shinkareva, S. V., Kim, J., Gao, C., & Wedell, D. H. (2020). Evaluative conditioning of affective valence. *Social Cognition*, *38*(2), 97-118. https://doi.org/10.1521/soco.2020.38.2.97

Wedell, D. H., & Senter, S. M. (1997). Looking and weighting in judgment and choice. *Organizational Behavior and Human Decision Processes*, *70*(1), 41-64. http://doi.org/10.1006/obhd.1997.2692

APPENDIX A

STUDY 1 SUPPLEMENTARY ANALYSES

A.1 TOTAL NUMBER OF FIXATIONS DURING INDUCTION

For participants who significantly preferred products paired with liked music (n = 9), a music valence (liked or disliked) × attribute valence (positive or negative) repeated-measures ANOVA was conducted on total number of fixations on the four attribute interest areas during induction. There was no main effect of music valence, $F(1, 8) = 3.277$, p = .108. There was a significant main effect of attribute valence, $F(1, 8) = 12.949$, p = .007, with negative attributes (M = 0.676) looked at significantly more often than positive attributes (M = 0.425). There was no significant music valence × attribute valence interaction, $F(1, 8) = 0.045$, p = .838.

For participants who significantly preferred products paired with disliked music (n = 4), a music valence (liked or disliked) × attribute valence (positive or negative) repeated-measures ANOVA was conducted on total number of fixations on the four attribute interest areas during induction. There was no main effect of music valence, $F(1, 3) = 0.944$, p = .403. There was a significant main effect of attribute valence, $F(1, 3) = 10.362$, p = .022, with negative attributes (M = 0.678) looked at significantly more often than positive attributes (M = 0.334). There was no significant music valence × attribute valence interaction, $F(1, 3) = 1.077$, p = .376.

For participants who did not show a significant effect in paired choice (n = 12), a music valence (liked or disliked) × attribute valence (positive or negative) repeated-measures ANOVA was conducted on total number of fixations on the four attribute interest areas during induction. There was no main effect of music valence, *F*(1, 11) = 1.210, *p* = .295. There was no main effect of attribute valence, *F*(1, 11) = 0.453, *p* = .515. There was no significant music valence × attribute valence interaction, *F*(1, 11) = 4.860, *p* = .050.

A.2 TOTAL FIXATION DURATION DURING INDUCTION

For participants who significantly preferred products paired with liked music (*n* = 9), a music valence (liked or disliked) × attribute valence (positive or negative) repeated-measures ANOVA was conducted on total fixation duration on the four attribute interest areas during induction. There was no main effect of music valence, *F*(1, 8) = 3.636, *p* = .093. There was a significant main effect of attribute valence, *F*(1, 8) = 10.895, *p* = .011, with negative attributes (*M* = 145.167 ms) looked at significantly longer than positive attributes (*M* = 88.623 ms). There was no significant music valence × attribute valence interaction, *F*(1, 8) = 0.785, *p* = .402.

For participants who significantly preferred products paired with disliked music (*n* = 4), a music valence (liked or disliked) × attribute valence (positive or negative) repeated-measures ANOVA was conducted on total fixation duration on the four attribute interest areas during induction. There was no main effect of music valence, *F*(1, 3) = 1.339, *p* = .331. There was no main effect of attribute valence, *F*(1, 3) = 1.662, *p* = .288. There was no significant music valence × attribute valence interaction, *F*(1, 3) = 0.043, *p* = .848.

For participants who did not show a significant effect in paired choice (*n* = 12), a music valence (liked or disliked) × attribute valence (positive or negative) repeated-measures ANOVA was conducted on total fixation duration on the four attribute interest areas during induction. There was no main effect of music valence, *F*(1, 11) = 1.049, *p* =

.328. There was no main effect of attribute valence, $F(1, 11) = 0.055$, $p = .819$. There was no significant music valence × attribute valence interaction, $F(1, 11) = 3.873$, $p = .075$.

A.3 TOTAL NUMBER OF FIXATIONS DURING PAIRED CHOICE

For participants who significantly preferred products paired with liked music (n = 9), a music valence (liked or disliked) × attribute valence (positive or negative) repeated-measures ANOVA was conducted on total number of fixations on the four attribute interest areas during paired choice. There was no main effect of music valence, $F(1, 6) = 0.291$, $p = .609$. There was a significant main effect of attribute valence, $F(1, 6) = 7.521$, $p = .034$, with negative attributes ($M = 1.004$) looked at significantly more often than positive attributes ($M = 0.911$). There was no significant music valence × attribute valence interaction, $F(1, 6) = 1.989$, $p = .208$.

For participants who significantly preferred products paired with disliked music (n = 4), there were insufficient participants with all design cells filled to test the effects of music valence and attribute valence on total number of fixations on the four attribute areas during paired choice.

For participants who did not show a significant effect in paired choice (n = 12), a music valence (liked or disliked) × attribute valence (positive or negative) repeated-measures ANOVA was conducted on total number of fixations on the four attribute interest areas during paired choice. There was a significant main effect of music valence, $F(1, 8) = 29.561$, $p = .001$, with attributes of products paired with liked music ($M = 1.647$) looked at significantly more often than attributes of products paired with disliked music ($M = 1.431$). There was no main effect of attribute valence, $F(1, 8) = 1.254$, $p = .295$. There was a significant music valence × attribute valence interaction, $F(1, 8) = 14.928$, $p = .005$, with negative attributes of positive products (*M* = 1.883) and positive attributes of negative products (*M* = 1.629) looked at significantly more often than positive attributes of positive products (*M* = 1.411) and negative attributes of negative products (*M* = 1.232).

A.4 TOTAL FIXATION DURATION DURING PAIRED CHOICE

For participants who significantly preferred products paired with liked music (n = 9), a music valence (liked or disliked) × attribute valence (positive or negative) repeated-measures ANOVA was conducted on total fixation duration on the four attribute interest areas during paired choice. There was no main effect of music valence, $F(1, 6) = 1.734$, $p = .236$. There was a significant main effect of attribute valence, $F(1, 6) = 5.545$, $p = .043$, with negative attributes (*M* = 199.065 ms) looked at significantly longer than positive attributes (*M* = 164.139 ms). There was no significant music valence × attribute valence interaction, $F(1, 6) = 2.135$, $p = .194$.

For participants who significantly preferred products paired with disliked music (n = 4), there were insufficient participants with all design cells filled to test the effects of music valence and attribute valence on total fixation duration on the four attribute areas during paired choice.

For participants who did not show a significant effect in paired choice (n = 12), a music valence (liked or disliked) × attribute valence (positive or negative) repeated-measures ANOVA was conducted on total fixation duration on the four attribute interest areas during paired choice. There was no main effect of music valence, $F(1, 8) = 2.234$, $p = .173$. There was no main effect of attribute valence, $F(1, 8) = 0.142$, $p = .717$. There was a significant music valence × attribute valence interaction, $F(1, 8) = 25.221$, $p = .001$, with negative attributes of positive products (*M* = 394.010) and positive attributes of negative products (*M* = 399.098) looked at significantly longer than positive attributes of positive products (*M* = 235.141) and negative attributes of negative products (*M* = 246.451).

A.5 REGRESSION ANALYSES PREDICTING FIXATIONS FROM CONTINGENCY AWARENESS

Four regression analyses were conducted for each participant group, predicting the average difference in total number of fixations or total fixation duration for congruent or incongruent attributes during induction or paired choice.

For participants who significantly preferred products paired with liked music (n = 9), contingency awareness was only a significant predictor of the average difference in total number of fixations during paired choice, R^2 = .556, $F(1,6)$ = 7.501, p = .034, standardized β = -0.745.

For participants who significantly preferred products paired with disliked music (n = 4), none of the regression analyses yielded significant results.

For participants who did not show a significant effect in paired choice (n = 12), none of the regression analyses yielded significant results.