Learning and Motivation 40 (2009) 74–93

Contents lists available at ScienceDirect

Learning and Motivation

journal homepage: www.elsevier.com/locate/l&m

The smart gut: Tracking affective associative learning with measures of “liking”, facial electromyography, and preferential looking

K. Carrie Armel [a,*], Carmen Pulido [b], John T. Wixted [c], Andrea A. Chiba [d,*]

[a] Stanford University School of Medicine, Stanford Prevention Research Center, Hoover Pavilion, N229,
211 Quarry Road, Mail Code 5705, Stanford, CA 94305-5705, USA
[b] San Diego State University, USA
[c] University of California, San Diego, USA
[d] Department of Cognitive Science, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0515, USA

ARTICLE INFO

Article history:
Received 19 February 2008
Revised 25 June 2008
Available online 14 September 2008

Keywords:
Human
Conditioning
Associative
Pictures
Visual stimuli
Food
Magnitude
Valence
Valence continuum
Liking rating
Preference
Physiological
Facial EMG
Preferential looking
Logistic
Learning parameters
Learning rate

ABSTRACT

We demonstrate here that initially neutral items can acquire *specific* value based on their associated outcomes, and that responses of physiological systems to such previously meaningless stimuli can rapidly reflect this associative history. Each participant participated in an associative learning task in which four neutral abstract pictures were each repeatedly paired with one of four foods that varied in valence and magnitude. Over the course of learning, participants’ “liking” ratings of and preferences for each picture came to reflect the value of the food with which it was paired. The abstract pictures also elicited physiological responses characteristic of the foods with which they were paired, including changes in facial electromyography (EMG) and preferential looking. A logistic modeling procedure showed that learning parameters, such as the rate at which participants learned the values associated with the pictures, were similar across food outcomes of different value.

---

* Corresponding authors. Fax: +1 650 723 6450 (K.C. Armel); +1 858 534 1128 (A.A. Chiba).
E-mail addresses: kcarmel@stanford.edu (K.C. Armel), chiba@cogsci.ucsd.edu (A.A. Chiba).

0023-9690/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.lmot.2008.06.003

You drive by a giant yellow M, the "Golden Arches". Inadvertently, your face tenses, your mouth gapes, and you think, "Yuck!" Your "gut" response to this icon is anticipating the greasy, nasty tasting food that once crossed your lips shortly after driving through the "Golden Arches". Seated next to you, your spouse continues to gaze at the "Golden Arches". Relaxed and smiling, your spouse can still taste the sumptuous burger that passed his lips shortly after driving through the Golden Arches. A giant yellow plastic M would seem to evoke little meaning by itself. Yet, billions of people have a personal history with this M, such that the M evokes a differential affective response for each individual.

Literally thousands of studies across tens of species have studied an associative learning or conditioning process whereby a neutral stimulus is paired with a biologically relevant outcome that facilitates a response. Subsequent to these pairings, the stimulus comes to independently elicit that response. Another consequence of pairing a stimulus with a biologically relevant outcome occurs when the previously neutral stimulus acquires the motivational value of the biologically relevant event. Numerous behavioral changes occur, once the stimulus has taken on the value of its outcome, including various "evaluative" responses to the cue (see Holland & Gallagher, 1999, for review). For example, a stimulus that has acquired positive value may be chosen over other stimuli (Chiba, Quinn, & Merzlyak, 2002), "liked" better than other stimuli (Johnsrude, Owen, Zhao, & White, 1999), may capture more attention than other stimuli (Derryberry, 1989), and can be used to reinforce learning (Hatfield, Han, Conley, Gallagher, & Holland, 1996). Accurately acquiring the value of the stimulus to reflect the current value of its associated outcome is adaptive for any species. This is not to imply that stimuli are only associated with biologically relevant outcomes, as stimuli can also acquire value through their associations with other relevant stimuli (Holland & Gallagher, 1999).

Many species adapt their facial expressions to a stimulus in accordance with the current motivational value of the stimulus (for a review, see Berridge, 2000). When viewing food items, most mammals exhibit one facial expression to aversive stimuli, and another to appetitive stimuli. These expressions are more pronounced if the stimuli are extremely appetitive or aversive. Further, such expressions appear to be specific to the motivational value of a stimulus rather than to the stimulus itself, as an animal's expression to the same stimulus may change based on the motivational value of that stimulus. For instance, rodents will respond with an appetitive reaction to a salty food when they are salt-deprived, but with an aversive reaction, to that same food, when they are satiated with salt. Characteristic taste reactivity or facial responses may also be elicited by the presence of a stimulus that was previously associated with an appetitive or aversive stimulus.

Facial expressions in humans are taken to reflect a person's evaluation of a stimulus. A popular theory of affect is based on the hypothesis that affect or emotion that is elicited by a stimulus varies along two different dimensions or continuums; valence (ranging from positive to negative) and arousal (ranging from low to high intensity) (Lang, 1995). Based on this theory, empirical evidence indicates that human facial reactions, as measured by electromyography (EMG), demonstrate predictable patterns to pictures of biologically relevant items that vary in valence and arousal (Bradley, 2000; Fridlund, Schwartz, & Fowler, 1984; Lang, Greenwald, Bradley, & Hamm, 1993; Schwartz, Brown, & Ahern, 1980). Specifically, corrugator (brow) muscles become progressively more relaxed as one moves from neutral to very pleasant pictures, and progressively more contracted when moving from neutral to very unpleasant pictures (Lang et al., 1993). In contrast, zygomatic (cheek) muscles become more contracted as the pictures become more positive or more negative, but are not contracted to neutral pictures (Lang et al., 1993). There is a tight correspondence between verbal evaluative judgments of the pictures and the specific facial expressions elicited while viewing those pictures.

In addition to stimuli considered to hold intrinsic biological relevance, after several pairings with an affective stimulus, a formerly neutral stimulus can invoke changes in a variety of human physiological responses; including alterations in heart rate, skin conductance, respiration, and eyeblink (Ayers & Powell, 2002). Despite demonstration that these physiological systems reflect stimulus value, whether or not human facial expression can adapt to reflect the current value of a stimulus remains controversial. A number of associative learning experiments have not demonstrated reliable facial EMG alterations to a stimulus following learning (Dimberg, 1987; Hamm, Greenwald, Bradley, & Lang, 1993; Sabatinelli, Bradley, Cuthbert, & Lang, 1996; Schienle, Stark, & Vaitl, 2001; Vaughan & Lanzetta, 1980). In contrast, after pairing a single stimulus with a single aversive outcome, two studies demonstrated that it is possible to elicit characteristically aversive facial EMG responses to the formerly neutral stimulus (Bunce, Bernat, Wong, & Shevrin, 1999; Flor, Birbaumer, Hermann, Ziegler, & Patrick, 2002). Thus, it appears that it may be possible for facial expression in humans to adapt to reflect the current motivational value of a single stimulus. It remains to be demonstrated whether or not human facial EMG can change to reflect current values of different stimuli across learning.

An advantage of studying associative learning processes in humans lies in the diversity of learned responses available for assessment by the experimenter. In addition to recording the patterns of alterations in a variety of physiological and behavioral responses to conditioned stimuli, the experimenter can ask the participant to evaluate a stimulus across learning. This technique is a common practice in the tradition of human evaluative conditioning and has been successfully utilized across a broad set of disciplines, ranging from learning psychology to consumer science (see De Houwer, Thomas, & Baeyens, 2001, for review). One of the original evaluative conditioning experiments demonstrated that systematic pairings of neutral pictures with pictures that are either liked or disliked resulted in participants' eventual report of corresponding liking and disliking of the originally neutral pictures (Levey & Martin, 1975). Following this original study, numerous researchers have replicated and refined techniques for achieving evaluative conditioning, and have developed additional measures of this conditioning (see De Houwer et al., 2001, for review). For example, preference judgments, or judgments about which stimulus is preferred out of several, have been utilized in some of these studies. In one study participants incidentally associated a food reward with abstract visual patterns during a working memory task. At the end of the experiment, participants had developed preferences for pictures that were frequently paired with reward over those which were not (Johnsrude et al., 1999). The ability to measure acquired motivational value using both alterations in "liking" and "preference" judgments provides reliability across different forms of behavioral expression.

The majority of studies of humans, non-human primates, and rodents examining the motivational value of a conditioned stimulus utilize a single positive and/or a single negative outcome (see Gallagher & Chiba, 1996, for review). Some studies successfully demonstrate that humans learn to associate multiple stimuli with outcomes of different valence (positive and negative) within a single experimental session. Yet these findings have not yet been extended to examine the association of multiple stimuli with outcomes that vary along a valence continuum (mildly good, very good, etc.) (Baeyens, Eelen, van den Bergh, & Crombez, 1989; Fulcher & Cocks, 1997). The ability to learn associations between locations or objects and outcomes that vary in magnitude or intensity across a valence continuum (e.g., mildly sweet vs. very sweet) has been assessed successfully in rodents (Chiba et al., 2002; Kesner & Williams, 1995; Peinado-Manzano, 1989; Pratt & Mizumori, 1998; Salinas, Packard, & McGaugh, 1993). Thus, it seems plausible that humans would also be able to associate multiple stimuli with outcomes that vary in magnitude, across a valence continuum.

In addition, it is not known whether the rate at which the initially neutral stimulus acquires value (or other learning parameters) varies depending upon the value of the outcome. Given that the current study examined acquisition during a simple evaluative conditioning experiment in which the visual conditioned stimuli (pictures) were pre-selected to be equivalent, and pictures were paired with their designated food outcome 100% of the time, one might expect that conditioned stimuli paired with the extreme food values (unconditioned stimuli) would more rapidly acquire associability (a preference or aversion). This expected result is based on the Rescorla–Wagner Model of Classical Conditioning (Rescorla & Wagner, 1972), in which the learning rate may be accelerated in cases where the unconditioned stimulus has higher associative strength (beta). This is likely to be represented as a steeper negative acceleration in the learning curves for the pictures paired with the extreme food values relative to those with more neutral food values. Here, it is important to keep in mind that this model is based on Pavlovian expressions of learning and associability. There is little data indicating whether a human preference rating might also reflect such a response. The present study attempts to address whether human participants are able to learn the motivational value of different pictures across multiple pairings of those pictures with foods that vary in magnitude across the valence continuum. The present investigation addresses this issue through the use of a novel associative learning task in which neutral abstract pictures are each repeatedly paired with one of four different foods, spanning positive and negative valence. Participants' repeated ''liking'' ratings of the pictures are used as an evaluative measure of learning, in keeping with previous experiments examining similar questions (Baeyens, Crombez, Hendrickx, & Eelen, 1995; Baeyens et al., 1989; Todrank, Byrnes, Wrzesniewski, & Rozin, 1995).

A logistic curve-fitting procedure is used to track participants' ''liking'' ratings of the pictures across learning, in order to determine the rate at which the pictures take on the motivational value of their associated foods. Although other studies have investigated learning in conditioning paradigms with humans (e.g., Baeyens, Eelen, van den Bergh, & Crombez, 1992; Lachnit, Lipp, & Gryschok, 2002), the curve-fitting procedure applied here allows for the quantification of various learning parameters, as well as an evaluation of the consistency of these parameters across stimuli of different value and across individuals. The logistic function also captures the dynamic nature of the learning process and is based on the entire data set, rather than on a subset of the scores (Christenfeld, Glynn, & Gerin, 2000). In addition to liking ratings, a second behavioral measure of the acquisition of learned value is acquired through participants' preferences for one picture over another, indicated by their choices on a forced-choice test.

A second experiment utilizes the same associative learning task, behavioral measures, and analyses, but additionally assesses the motivational value of visual stimuli by measuring facial EMG and preferential looking. Following several repetitions of the associative learning task, patterns of corrugator and zygomatic muscle responses to the conditioned pictures were recorded (Lang et al., 1993). By measuring patterns of facial expression to the pictures, the study addresses whether human facial expressions adapt to reflect the current value of a stimulus. Analysis of each participants' preferential looking, or time spent looking at one picture relative to another, was used as a convergent measure of motivational learning.

Experiment 1

Method

Participants

A total of 48, University of California at San Diego students between the ages of 18 and 25 enrolled in the experiment through the psychology participant pool. All had normal or corrected to normal vision and spoke English. Participants were screened and eliminated if they used drugs regularly, or had a history of eating, psychiatric, or neurological disorders. Upon completing the experiment, participants received experimental credit to fulfill a course requirement.

Materials

Picture selection

Through pilot work, four pictures were selected for the experiment from a set of twenty abstract black and white designs that could not easily be verbally labeled. Pictures were selected if they were equally discriminable (so that participants would not confuse pictures that were aesthetically more similar) and neutral. To score the pictures on these dimensions, undergraduates gave same-different judgments, and also provided ''liking'' ratings (evaluated on a 1–7 Likert scale from dislike very much to like very much, where a 4 was neutral). We analyzed the same-different data using a multidimensional scaling (MDS) procedure (Sergent & Takane, 1987; Takane & Sergent, 1983). This allowed us to use correctness of the same-different judgments, in conjunction with reaction time to make those judgments (reaction time is faster when pictures are more different), to determine the distances in aesthetic space between the pictures. Of the eight pictures determined to be equally discriminable, the four with liking rating means closest to neutral were used in the experiments reported here, and these can be viewed in Fig. 1.

Fig. 1. The four pictures used in Experiments 1 and 2. The pattern on the far right was the mask. Each picture was paired with one food, and the pairings remained constant throughout the experiment for each participant. The specific picture–food pairings were randomly generated for each new participant.

Food selection

During pilot work, 40 participants provided liking ratings for over 25 foods. Participants tasted a given food and evaluated it on a 1–7 Likert scale from dislike very much to like very much. After each taste, participants took a sip of chilled bottled water to cleanse their palate for the next food. All foods were FDA approved and available at standard supermarkets, and the set-up was hygienic (using all disposable items). Foods were selected from this pilot work for the experiments presented here if they were rated almost unanimously by participants along the liking scale (i.e., valence continuum) around 1 (very negative), 3 (slightly negative), 5 (slightly positive), and 7 (very positive). The final foods included: A pinch of baking powder served as the "yuck". A pinch of cornmeal served as the "bad" food. A pinch (approximately 15 granules) of cane sugar served as the "goody". A mouthful of ice cream served as the "delicious" food (participants got their choice of Breyer's chocolate, vanilla, or strawberry, or frozen soymilk if they had dairy restrictions).

Design and procedure

For the present study participants were seated in a comfortable chair at a table facing a computer screen. To their left was the seated experimenter, the food, and a vertical partition placed on the table that blocked their view of the food (see Fig. 2). The experimenter wore headphones so that she could receive instructions from the computer program regarding food identity. At the base of the partition on the participant's side was a port where spoons were placed by the experimenter. The top of the port was opaque and thus blocked the participant's view of the food on the spoon, but allowed them to see the protruding handle to grasp the spoon. To further prevent participants from seeing food on the spoon, participants wore a special pair of glasses with an opaque white paper attached from arm to arm along the base of the glasses and protruding forward from their face by about three inches. Thus, participants' lateral and inferior peripheral vision was occluded. Participants were instructed that to taste any food during the course of the experiment, they should look down at the spoon handle, grasp the spoon (at this point, the glasses block the view of the spoon), look at the computer screen, and then lift the spoon and taste the food.

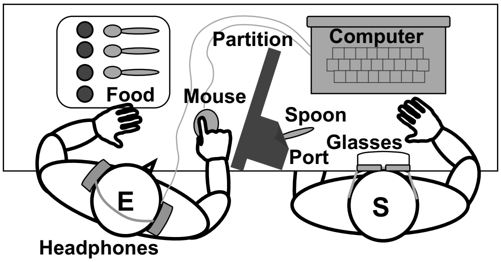

Fig. 2. The set-up for Experiment 1. The set-up for Experiment 2 was identical except that it additionally included electromyographic (EMG) equipment and a video camera for recording eye movements. E, experimenter; S, participant.

Next, food restrictions and preferences were investigated. It was verified that participants did not have food restrictions or allergies to any food used in the experiment. It was also verified that participants, as instructed upon enrolling in the experiment, did not eat for two hours prior to the experiment. This was to control for hunger/food motivation across participants. To determine each participant's food preferences, they were given a taste test (described under the "Food Selection" section above, except that they only tasted the four foods that were selected through the pilot work). For the rest of the experiment, instructions appeared on the computer screen and the experimenter read these aloud.

There were a total of ten experimental blocks, each of which was divided into three phases: the Training, the Liking Test, and the Preference Test (see Fig. 3). Instructions were first given for the Training trials, in which a participant saw a neutral abstract picture and then tasted a food of a particular outcome value. There were a total of four pictures, and also four foods that were of outcome values varying from "dislike very much" to "like very much". Each picture was paired with only one of the foods, and each food to only one picture, and these one-to-one pairings remained constant for the entire experiment. Participants were told: "The experimenter will place a spoon on the table. You should: (1) Pick up the food when the experimenter says "okay". (2) Look at the computer screen. (3) Taste the food." The instructions continued, "A picture will appear on the screen shortly before you taste the food in your mouth. Make sure you look at this picture. You should not worry about timing the food with the picture." The experimenter was in control of the timing, and initiated the 0.6 s

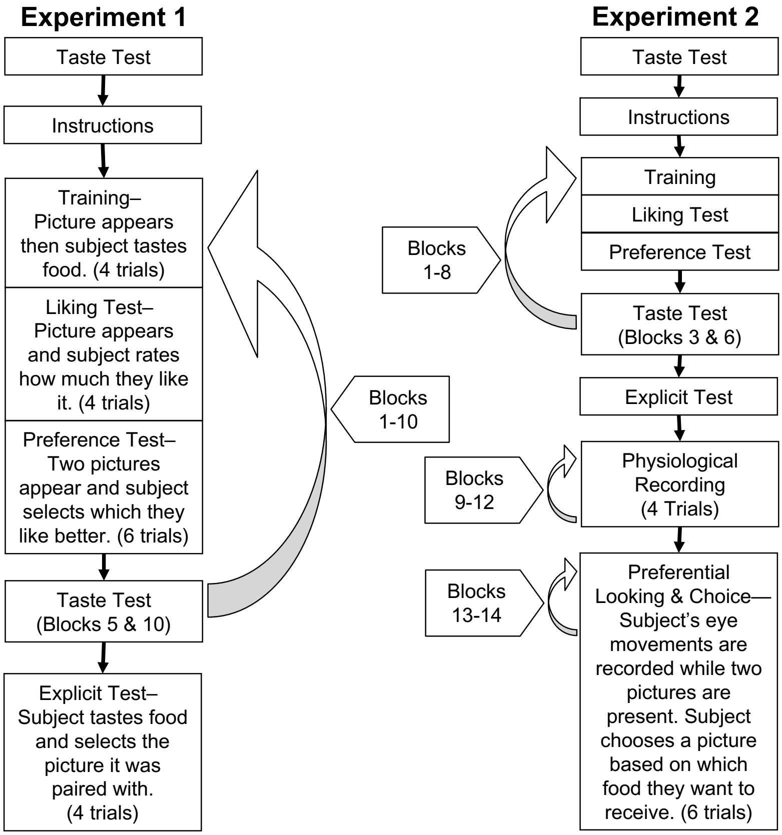

Fig. 3. Design for Experiments 1 and 2, as well as descriptions of each trial type. Trials within each square were randomized so that trial order was different across participants and also across blocks for a given participant.

picture presentation by clicking a computer mouse on her side of the partition when the food was approximately halfway between the food port and participant's mouth. This ensured that the participant tasted the food after the picture appeared on the screen but before it disappeared. The next screen read, "After each picture disappears, a pattern will appear, making the picture difficult to remember. You do not need to remember the picture—this is *not* a memory experiment." The pattern was a mask presented for 0.48 seconds that controlled for participants' variability in visual memory by preventing all participants from maintaining an after-image of the picture. Participants were instructed to give the spoon back to the experimenter and take a sip of water after the screen went blank. Participants then had three practice trials of viewing a picture and tasting a food (a small amount of cornstarch was used as a "neutral" food in these practice trials).

Participants then received instructions for the test portions of the experiment. The Liking and Preference tests determined whether a participant learned the associations between the pictures and their corresponding outcome value. On the Liking Test trials, a participant saw a picture and was asked how much they liked it on a scale of 1–7. A rating of one means they disliked it very much, a rating of four was neutral, and a rating of seven indicated that they liked it very much. Participants were familiarized with the rating scale and were also told, "Each picture disappears quickly and is followed by a pattern [the mask], making it hard to remember. We do this because we want you to go with your "gut" feeling as to how much you like the picture. You shouldn't think too hard or try comparing the picture to anything else. Just view each picture on its own and judge how much you like it." Three Liking Test practice trials ensued. We anticipated that over time each picture would take on the value of the food paired with it, so that a participant would give a picture paired with Yuck a very low "liking" rating (e.g., 1), one paired with Bad a rating slightly below neutral, one paired with Goody a rating slightly above neutral, and one paired with Delicious a very high "liking" rating (e.g., 7).

Then participants received instructions for the Preference Test, in which they selected which picture they preferred out of two that appeared simultaneously. Participants were instructed: "*Remember, go with your gut intuition*." Participants practiced making preference judgments for three trials. We anticipated that for each pair, a participant would "prefer" the picture paired with the more appetitive (or less repelling) food.

Participants were told they would cycle through the three parts of the experiment several times (see Fig. 3). There were a total of ten experimental blocks that ensued, each of which was divided into three phases of trials described above: the Training, the Liking Test, and the Preference Test. A participant received each of the four picture–food pairings once in each Training block, and their presentation order was randomized across blocks. The specific picture–food pairings were randomized across participants, such that the pictures were paired with different food values for different participants. In each Liking Test block there were four trials, one for each picture, and the order of the trials was randomized across blocks. In each Preference Test block, there were six trials, so that a participant was queried about every possible pair of pictures, and the order of the trials was randomized across blocks. The 10 blocks were punctuated only by taste tests after blocks 5 and 10, to determine whether taste preferences were stable over the course of the experiment.

After the last taste test, participants' visual perception and memory abilities were assessed with the Explicit Test of the picture–food pairs. For this test, participants were presented with the four pictures they saw during the task. They were given each of the foods, one at a time, and asked to identify the picture that had been paired with each food during the experiment. They were also asked to give a confidence rating, from 1 to 7 (one being completely guessing and seven absolutely certain), as to how certain they were that the picture was paired with that food.

Data analysis

Liking ratings: Logistic regression

To determine whether each picture acquired the motivational value of its associated food and to determine whether there exist differences in characteristic functions between pictures paired with different foods, we fit a version of the logistic function. This descriptive function follows a sigmoidal path that captures the expected gradual transition from an initial, neutral baseline liking value for a picture to its final liking value after multiple training trials. The function is:

$$Y(x)_c = b + A_c/(1 + \exp(-(x - m)/s))$$

Here, $Y(x)_c$ is the regressed "liking" rating (or picture value) for a particular condition (Delicious, Goody, etc.) on block x, where x ranges from 1 to 10. The four parameters include: b which is the baseline or Y intercept when $x = 0$ (i.e., the neutral "liking" rating for all pictures before the experiment begins, which should be about 4 on the 1–7 scale), A_c which is the asymptotic change from baseline following training (roughly the final amount of "liking" added to or subtracted from pictures paired with each food), s which is the rate of transition between the pre- and post-learning asymptotes (larger values indicate a relatively slow transition; smaller values indicate a relatively fast transition), and m which is the midpoint of the transition between the pre- and post-learning asymptotes measured by block (equivalent to the time halfway through learning). Only one baseline value was used for all four conditions because the picture–food pairings were randomized across participants; thus, the baselines before block 1 for pictures paired with different foods could not differ.

Learning midpoint and rate (as indicated by m and s, respectively) seemed to be consistent across picture–food pairings. As a result we used one parameter for rate and one for midpoint in our model. This was determined by testing whether one parameter would suffice, or whether four (one for each picture–food pairing condition) would significantly improve the fit of the model to our data. We addressed this question for midpoint, slope, and asymptotes independently, and will use midpoint to illustrate the procedure. After averaging the data across participants, we used the previously described equation with one midpoint parameter to obtain initial estimates for all the parameters. These estimates were then held constant in a new equation where four midpoint parameters were allowed to vary. We then determined whether four midpoint parameters significantly improved the variance accounted for by the model—put another way, whether four significantly minimized the residual sum of squares. Within the limits of measurement error, we were unable to detect significant differences in the variance accounted for by the two models. Because there was no advantage to allowing midpoint to vary across conditions, we used only one midpoint parameter in the remaining analyses. The same procedure was used and conclusions drawn for slope. However, asymptotic fit was significantly improved by allowing it to vary across conditions. Thus we used the equation specified above for the remaining analyses. The logistic function was fit to 40 data points (10 liking ratings from each of four conditions), and the estimated parameters were b, m, s, A_1, A_2, A_3, and A_4.

To verify the estimates produced by the model, we then computed estimates for each individual participant. We report the number of participants the model was able to fit, as well as a new set of estimates derived from this subset of participants. Excluding the few participants the model was not able to fit may provide a better indication of true parameter estimates.[1] With the exception of this logistic procedure on individual participants' data, all analyses reported here included all participants.

Liking ratings and preferences: ANOVA. To determine whether the "liking" ratings of each picture differed from the others and whether changes in the value of the pictures changed over time, A two-way (picture value) × (block cluster) repeated measures ANOVA was performed. Step-down ANOVAs were used to analyze each main effect. Planned comparisons of sets of pictures were accomplished using contrasts. Post hoc comparisons were accomplished using Student–Newman–Kuels tests.

A confidence level of 95% ($\alpha = .05$) was used for all statistical tests unless otherwise noted. While the logistic regression was performed on data from all 10 recorded time-points, the remaining behavioral analyses for Experiment 1 were performed on data at five time-points or "block clusters". That is, "liking" ratings were averaged for blocks 1 and 2, 3 and 4, 5 and 6, 7 and 8, and 9 and 10 prior to analyses to decrease variability in each participant's data. Preference judgments were similarly averaged.

---

[1] Specifically, participants were excluded if their parameter estimates were not valid—for example, the asymptotic estimate for one of the picture values was outside our 7-point scale. Typically, for these participants, the model also produced parameter estimates with high standard errors, and it accounted for a low level of variance in the participant's data. Examination of the raw data for these participants revealed aberrant data patterns (they did not show learning, their learning began at block 1 and did not improve, or they exhibited learning and then stopped showing it).

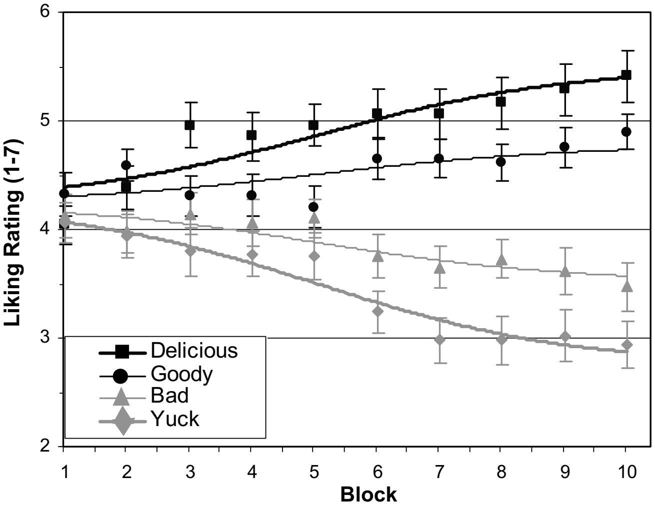

Fig. 4. Picture liking ratings for Experiment 1. Pictures paired with each taste across all blocks are displayed. Pictured here are both the logistic regression curves and actual means with standard errors (1 *SEM*).

Results

Liking ratings: Logistic regression—group data

Applying the described logistic regression model to the group "liking" averages generated numerous best-fitting parameter estimates (the "liking" ratings and fitted functions are depicted in Fig. 4). The baseline parameter confirmed that the pictures did begin with a neutral motivational value. The baseline, in terms of the 1–7 "liking" scale, was 4.25 (ASE (asymptotic standard error) = .05). This value was approximately neutral on the "liking" scale.

Participants were observed to be halfway through the learning process on the fifth block of learning trials. This was demonstrated by a midpoint parameter expressed in terms of block as 5.07 (ASE = .68). The model indicated that participants learned the associations at a fairly rapid rate. The slope, which measures the rate of change between baseline and asymptote (a larger value indicates slower learning), was 2.03 (ASE = .47).

Once participants' "liking" ratings of each picture reached asymptote, their ratings (based on change from baseline) of the pictures decreased monotonically from Delicious to Yuck. The asymptotic values relative to the baseline were as follows: Delicious 1.25 (ASE = .18), Goody 0.54 (ASE = .13), Bad −0.73 (ASE = .14), and Yuck −1.49 (ASE = .20). Thus, the pictures acquired the predicted motivational values, as seen by the correct rank-order of the initially neutral pictures according to their outcomes. The variance accounted for by our logistic function and these parameter estimates was high at 95.5%.

Liking ratings: Logistic regression—individual data

To verify the estimates produced by the model, we computed estimates for each individual participant. In fitting individual curves, the model was able to fit 35 of 48 participants[2]. The averages of the individual estimates that the model was able to fit were as follows: the baseline was 4.30 (*SE* = .14), midpoint was 3.63 (*SE* = .33), slope was 0.72 (*SE* = .14), the asymptote for Delicious was 1.11 (*SE* = .31), Goody was 0.34 (*SE* = .23), Bad was −0.84 (*SE* = .24), and Yuck was −1.79 (*SE* = .24). The average variance accounted for by our logistic function and the parameter estimates was 62.1% (*SE* = 2.8%). As one might expect, the exact estimates changed because of excluding some participants. In general the individual participant data reflect a more rapid rate of learning.

---

[2] Referring to pyramidal system including the motor strip just anterior to the central sulcus, the internal capsule, and brain stem and spinal cord motor nuclei.

When these same participants' "liking" ratings were analyzed (based on asymptotic estimates), it was shown that the pictures came to reflect the motivational value of their associated food. To determine whether the four pictures acquired distinct values for the participants, a one-way repeated measures ANOVA for the asymptotic picture values was performed. The ANOVA was significant (all $F(3,34) = 26.95$, $p < .0001$). To determine whether all the pictures acquired distinct value, we performed Student–Newman–Keuls post hoc comparisons. The six comparisons of each picture value to every other were all significant, indicating all pictures acquired distinct value.

Liking test ANOVA

The results of the logistic regression were confirmed using a two-way (picture value) × (block cluster) repeated measures ANOVA. The ANOVA demonstrated that pictures differed in their value, and also that they changed in value over time. Main effects for picture value ($F(3,47) = 20.50$, $p < .0001$), block cluster ($F(4,47) = 4.00$, $p < .01$), and their interaction ($F(12,188) = 10.44$, $p < .0001$) were all significant.

To determine whether the four pictures acquired distinct values for the participants, and at what point in time this occurred, step-down repeated measure ANOVAs for picture value were computed at each block cluster. The pictures did acquire different values given that the ANOVAs for the last four block clusters were significant (all $F(3,47) > 6.51$, $p < .001$). To determine whether all the pictures acquired distinct value, we performed Student–Newman–Keuls post hoc comparisons between each of the picture values for the final block cluster, resulting in six comparisons. We analyzed only the final block cluster to minimize the number of comparisons, and because participants would have developed the strongest picture–food associations at that point. All across-valence comparisons were significant (Delicious vs. Bad, Delicious vs. Yuck, Goody vs. Bad, and Goody vs. Yuck). For pairs of the same valence but different magnitude (Delicious and Goody, and also Yuck and Bad), liking ratings showed the predicted relative relationships, as seen in Fig. 4, although these comparisons were not significant.

Values were learned over time. To determine whether the picture values were acquired over time, step-down repeated measures ANOVAs for each picture value were computed across the block clusters. ANOVAS for Delicious, Bad, and Yuck were all significant at $p < .0001$ ($F(4,47) > 7.76$). Goody was not significant, probably because participants gave it a higher "liking" rating at the outset and maintained that high rating (see Fig. 4). This suggests that, except for Goody (which was learned immediately), the rest of the pictures acquired their value from the food pairing after the first block cluster, a pattern that can be seen in Fig. 4.

Preference test

A two-way ANOVA (block cluster) × (picture pair) repeated measures ANOVA showed that over the course of the experiment participants came to prefer the picture of greater value when presented with a pair of pictures. According to our coding, a "correct" response occurred when participants could discriminate between the two pictures and selected the one of greater value. Overall, participants did not exhibit this preference more strongly for particular picture pairs, indicating that participants showed within, as well as across, valence preferences. Specifically, there was a main effect for block cluster ($F(4,46) = 19.24$, $p < .0001$) but not picture pair. There was a significant interaction between block cluster and picture pair ($F(20,46) = 1.59$, $p < .05$), suggesting that some of the pairs may be discriminated with less learning than others.

To determine when participants began to exhibit a preference for the pictures of greater value, step-down repeated measures ANOVAs were computed for each block cluster. ANOVAs for the first two block clusters were not significant, but those for the last three block clusters were (all at $F(5,47) > 2.36$, $p < .05$). This suggests that, based on the Preference Test, values for the pictures were learned by the third block.

Explicit test

Participants correctly paired the pictures with the foods at the end of the experiment. In a one-sample *t*-test with the hypothesized mean equal to chance, or 0.25 (participants selected one of four pictures), participants were significantly better than chance at correctly selecting the picture that went with a particular food for all four picture–food pairings ($t(46) > 4.78$, $p < .0001$). Furthermore, "liking" ratings on block 10 were correlated with performance on the explicit tests for positive pictures, and inversely related for negative pictures (Delicious $r(47) = .42$, $p < .01$; Goody $r(47) = .40$, $p < .01$; Bad $r(47) = -.31$, $p < .05$; Yuck $r(47) = -.537$, $p < .0001$). This shows that for each of the picture–food pairings, when a participant's ''liking'' rating for a picture was consistent with its outcome value, performance on the explicit test was better.

Taste test

Taste ratings for each food remained relatively stable across the three taste tests, as indicated by non-significant findings on a one-way repeated measures ANOVA.

Discussion

The results of this experiment indicate that initially neutral pictures acquire the motivational value of their associated foods. This acquisition results from repeated pairings of the pictures with their associated foods. These results are evident in Fig. 4, when examining the correctly rank-ordered logistic regression asymptotes. The final ''liking'' rating for each picture was different than that of every other, as supported by a post hoc comparison of the logistic regression analysis. Post hoc analysis of the raw means of the picture ratings, however, indicated that the two within valence comparisons were not significant. This suggests that ''liking'' ratings for Delicious and Goody, and also for Yuck and Bad, were only marginally distinct. This could be due to the limitations of the rating paradigm, as a 7-point scale is compact and does not provide much room for variations in rating.

Liking rating data fits a logistic function, suggesting that there is a learning period during which time ''liking'' ratings for the initially neutral pictures are changing, as well as pre- and post-learning asymptotic periods during which little learning is occurring. According to the logistic regression on group data, participants were observed to be halfway through the learning process on the fifth block of learning trials. The latency and rate of this learning was consistent across the various food values (as determined by no systematic or significant trend differences for midpoint or rate when the asymptotes were fixed at their original estimates). One possible exception to this was for Goody. A non-significant step-down ANOVA for picture value across block cluster revealed that Goody did not change value over time, perhaps because its value was learned immediately. This may be due to the innately reinforcing value of sugar, or that sugar had an initially higher outcome value than the other foods regardless of likability (e.g., because of the concentration of sweetness). This is an interesting issue to attempt to dissociate in future experiments.

The observation that the latency and rate of learning were relatively consistent across the different food values (unconditioned stimulus values) was counter to the prediction set forth by the Rescorla–Wagner Classical Conditioning Model. Despite the inherent differences in the associative strength of the various foods, these differences were not sufficient to drive learning rates in a divergent manner. Perhaps this is due to the fact that the primary learning measure in the present study is a retrospective measure rather than a somatic measure. Future experiments can take advantage of the obtained changes in facial expression in order to see whether this measure reflects learning rates that are commensurate with the predictions made by the Rescorla–Wagner model. Additional parametric manipulations will also allow the present experimental approach to address further questions regarding basic learning theory and the neuroanatomical basis of various aspects of evaluative conditioning.

On the Preference Task participants showed learning of the picture–food associations by the third block, which is consistent with the Liking data. Participants discriminated equally well between the different picture pairs.

Experiment 2

Method

Participants

Twenty-five students participated who were of the same demographics and population as in Experiment 1. Upon completing the experiment, participants received experimental credit to fulfill a course requirement.

Design and procedure

The design and procedure were similar to those in Experiment 1. In Experiment 2, however, there were eight instead of 10 experimental blocks. Participants received a taste test before beginning the experiment, and also after blocks 4 and 6.

Following the eighth block of learning, participants received four Physiological blocks, in which facial EMG was recorded. These blocks were similar to the Training blocks except that picture presentation lasted for six seconds and food tasting was not initiated until mask offset. The physiological measures included facial (EMG) of the corrugator and zygomatic muscles.

After the Physiological blocks, there were two blocks of Preferential Looking and Preference (forced-choice) trials. These trials were similar to the Preference Test trials in that two pictures appeared on the screen at the same time. The pictures were presented for four seconds and participants were instructed to look at them the entire time they were on the screen. When the pictures disappeared, participants were asked to "choose one" of the pictures to receive the corresponding food. Participants were videotaped during this procedure and eye movements were analyzed off-line. Each block consisted of the six possible picture pairs presented in random order.

In Experiment 2 we changed the Goody and Bad foods to increase the magnitude differences between them and the Delicious and Yuck foods. Thus, Goody was changed from pure sugar to sugar cut with cornstarch, and Bad was changed from cornmeal to flour.

For a summary of the experimental design, see Fig. 3.

Physiological measures

Facial electromyography. We recorded facial electromyographic (EMG) activity over the left corrugator and zygomatic sites to measure participants' perceived valence of the pictures. Electrodes were placed according to methods established by Lang et al. (1993). Facial EMG is a particularly appealing measure of learning in this experiment given that, while "liking" ratings and preference judgments may be susceptible to task demands, facial EMG is not. This is because participants were naïve as to the purpose of the face electrodes (that they were measuring affective learning) and to the specific relationships between facial muscle movements and valence.

Facial EMG was recorded with silver/silver-chloride 4 mm diameter electrodes and routed through Biopac's MP100 series bioamplifiers. Signals were bandpass filtered from 90 to 250 Hz, integrated (averaging over 200 samples/second), and rectified. Change scores were calculated separately for EMG activity at each facial muscle site by subtracting the mean activity during the one second preceding picture onset (baseline) from the average response during the 6-s picture viewing interval. If the one second preceding picture onset was noisy, a more stable baseline was selected as close to (but preceding) picture onset as possible.

Preferential looking. Videotapes of participants' eye movements were analyzed to determine how long participants looked at each side of the screen. Preferential looking procedures, videotape coding, and analyses were similar to those used in preferential looking studies with infant humans (Choi, McDonough, Bowerman, & Mandler, 1999; Schafer & Plunkett, 1998). Video was recorded with a Sony Digital 8 Camcorder at 30 frames per second, with the video camera located behind the laptop screen on which participants viewed stimuli. MGI VideoWave software allowed coders to view the video off-line and advance it one frame at a time to determine the duration that participants' eyes were fixated within a given half of the laptop screen. For each trial, total time that a participant viewed each half of the screen was determined, and the total viewing duration of the shorter side was subtracted from that of the longer side to produce a difference score for each trial. Because there were two blocks of preferential looking trials, each one including the six different picture pairs, the difference scores for the two trials of the same pair were averaged to produce six preferential looking difference scores for each participant. Time participants spent blinking or looking away from the pictures (e.g., between them while making a saccade from one picture to the other) was excluded from analyses. Coders were blind as to which side of the screen displayed the picture of "greater value".

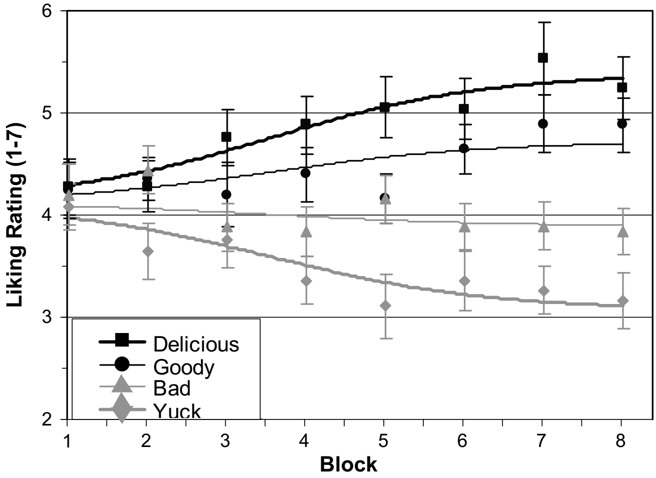

Fig. 5. Picture liking ratings for Experiment 2. Pictures paired with each taste across all blocks are displayed. Pictured here are both the logistic regression curves and actual means with standard errors (1 *SEM*).

Results

Behavioral results

The behavioral results reflected the same behavioral patterns of learning as observed in Experiment 1.

Liking ratings: Logistic regression—group data. Applying the described logistic regression model to the group "liking" averages generated numerous best-fitting parameter estimates (the "liking" ratings and fitted functions are depicted in Fig. 5). Again, the pictures were determined to be neutral prior to learning. The baseline for picture "liking" was 4.12 (ASE = .08), which was approximately neutral on the "liking" scale.

Learning proceeded rapidly, as participants were halfway through learning between blocks three and four and the slope was 1.38 (with smaller numbers indicating faster learning). The midpoint of the learning curve was 3.53 (ASE = .55), suggesting participants were halfway through learning before block 4. The slope was 1.38 (ASE = .46), thus learning was slightly faster than in Experiment 1.

Once participants' "liking" ratings of each picture reached asymptote, their ratings (based on change from baseline) of the pictures decreased monotonically from Delicious to Yuck. The asymptotic values relative to the baseline were as follows: Delicious 1.27 (ASE = .20), Goody 0.60 (ASE = .15), Bad −0.22 (ASE = .14), and Yuck −1.05 (ASE = .17). As predicted, the pictures acquired the motivational values of their associated foods, as evidenced by the correct rank-order of the values (from Delicious to Yuck). The variance accounted for by our logistic function and these parameter estimates was high, 93.0%.

Liking ratings: Logistic regression—individual data. To verify the estimates produced by the model, we computed estimates for each individual participant. In fitting individual curves, there were 19 of 25 participants that the model was able to fit[3]. As in Experiment 1, analysis of the individual data that the model was able to fit confirmed the overall results of the group data. The baseline was 3.90 (*SE* = .26), again close to neutral. Learning proceeded swiftly, as evidenced by a midpoint of 2.76 (*SE* = .29) and a slope of 0.79 (*SE* = .33). The asymptote for Delicious was 1.53 (*SE* = .37), Goody was

---

[3] Referring to the extrapyramidal system which is composed of a group of highly interactive neural circuits, including areas of frontal cortex and many subcortical nuclei, each of which contributes its own specialized influences to the final motor response.

0.69 ($SE = .35$), Bad was 0.21 ($SE = .29$), and Yuck was −0.86 ($SE = .38$). Thus, the eventual picture ratings reflected the value of their associated foods. The average variance accounted for by our logistic function and the parameter estimates was 57.0% ($SE = 5.0\%$). As one might expect, the exact estimates changed because of excluding some participants.

The pictures came to hold different motivational value after learning had occurred. To determine whether the four pictures acquired distinct values for the participants, a one-way repeated measure ANOVA for the asymptotic picture values was performed. Six of the 24 participants were excluded from analyses because the model could not be fit to their data, based on the same criterion used in Experiment 1. For the remaining 19 participants, the ANOVA was significant ($F(3,18) = 13.86$, $p < .0001$). To determine whether all the pictures acquired distinct value, we performed Student–Newman–Keuls post hoc comparisons. The six comparisons of each picture value to every other were all significant, except for the comparison of Goody to Bad.

Liking test. An overall two-way repeated measures ANOVA (picture value) × (block cluster) showed that pictures differed in their value. The main effect for picture value ($F(3,23) = 10.98$, $p < .0001$) was significant, confirming that the ratings of the pictures had changed. A reliable interaction between picture value and block cluster indicated that pictures changed value differentially across learning ($F(9,69) = 5.00$, $p < .0001$).

Again, pictures that were associated with extreme points on the valence continuum were significantly different from one another. Step-down repeated measures ANOVAs for each picture value were computed, and these were significant for Delicious ($F(3,23) = 6.80$, $p < .001$) and Yuck ($F(3,23) = 3.33$, $p < .05$), but not for Goody or Bad.

Preference test. An overall two-way repeated measures ANOVA (block cluster) × (picture pair) showed that over the course of the experiment participants came to prefer the picture of greater value, when presented with a pair of pictures. The main effect for block cluster ($F(3,23) = 8.93$, $p < .0001$) was significant, but not for picture pair or their interaction. This suggests that participants exhibited preferences for items of greater value across all picture pairs. No step-down ANOVAs for block cluster were significant.

Explicit test. Participants correctly paired the pictures with the foods at the end of the experiment. In a one-sample *t*-test with the hypothesized mean equal to chance, or 0.25, participants were significantly better than chance at correctly selecting the picture that went with a particular food for all four picture–food pairings ($t(23) = >2.81$, $p < =.01$). Furthermore, “liking” ratings on block 10 were correlated with performance on the explicit tests for positive pictures, and inversely related for Yuck (Delicious $r(24) = .54$, $p < .01$; Goody $r(24) = .59$, $p < .01$; Bad not significant; Yuck $r(24) = -.46$, $p < .05$). This shows that for each of the picture–food pairings, when a participant’s “liking” rating for a picture was consistent with its outcome value, performance on the explicit test was better.

Taste test. Again, the taste ratings for each food remained relatively stable across the taste tests, as demonstrated by a non-significant one-way ANOVA.

Physiological results

Facial electromyography. Participants exhibited the predicted linear relationship between picture pleasantness and corrugator response, and the quadratic relationship between picture pleasantness and zygomatic response. Four blocks of facial EMG data were averaged, producing a total of eight data points per participant—one for each picture value for the left corrugator muscle, and one for each for the left zygomatic. A simple regression for corrugator EMG was significant ($r(1,87) = .42$, $p < .0001$). That is, corrugator activity increased in response to pictures paired with more aversive foods, indicating perceived unpleasantness as established by previous work (Bradley, 2000; Fridlund et al., 1984; Lang et al., 1993; Schwartz et al., 1980). In contrast, corrugator activity decreased in response to pictures paired with more pleasant foods, indicating perceived pleasantness. This provides support that the pictures acquired the hedonic values with which they were paired. A second-order polynomial regression for zygomatic EMG was significant ($r(2,79) = .32$, $p < .05$). That is, zygomatic activity increased in response to pictures paired with increasingly pleasant foods, but also increased in response to pictures paired with very aversive foods, indicating perceived pleasantness and unpleasantness respectively. Again, this provides support that the pictures acquired the hedonic values with which they were paired. Figs. 6 and 7 illustrate these patterns.

Preferential looking. Participants looked longer at pictures of greater value. One-sample one-tailed *t*-tests were employed, and the statistic was computed by comparing the difference score for each picture pair with that of chance, a difference score of zero. Participants looked longer at the picture of greater value for all picture combinations ($t(15) = >1.91$, $p < .05$) except Delicious vs. Goody see Fig. 8. This is an additional piece of evidence that individuals preferred the better item in each pair, based on previous work showing that individuals spend more time looking at items that they prefer (Shimojo et al., 2003; Karjbich, Armel, & Rangel, in preparation).

Discussion

Behavioral findings

Again, the initially neutral pictures acquired the motivational value of their associated foods. This was achieved following repeated pairings of the pictures with the foods and can be observed in Fig. 5 in the correctly rank-ordered logistic regression asymptotes. The fact that the same basic pattern of results occurred across two separate groups of participants participating in two different experiments lends great reliability to the associative learning task. In contrast to the previous experiment, the liking rating values for Goody and Bad remained fairly close to neutral. This might be due to the fact that we modified the foods used for these picture values from Experiment 1 so that they would be “farther away” in “liking” ratings from Delicious and Yuck on our compact 7-point scale.

As in Experiment 1, “liking” rating data fit a logistic regression function, and the latency and rate of pairing the food values to the pictures was consistent across the various values. Values were acquired slightly earlier (the midpoint of learning was between blocks two and three) and more rapidly in Experiment 2, relative to Experiment 1. Confirming the results of Experiment 1, in the Preference Task participants discriminated equally well between the different picture pairs.

Physiological findings

The present experiment demonstrates that facial muscle responses can be elicited by previously neutral pictures that acquire motivational significance within a single learning session. The partici-

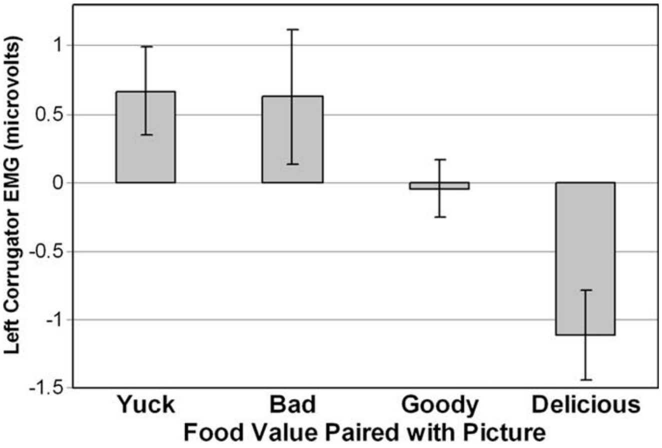

Fig. 6. Left corrugator EMG responses to pictures paired with different food values for Experiment 2. Standard error bars represent 1 *SEM*.

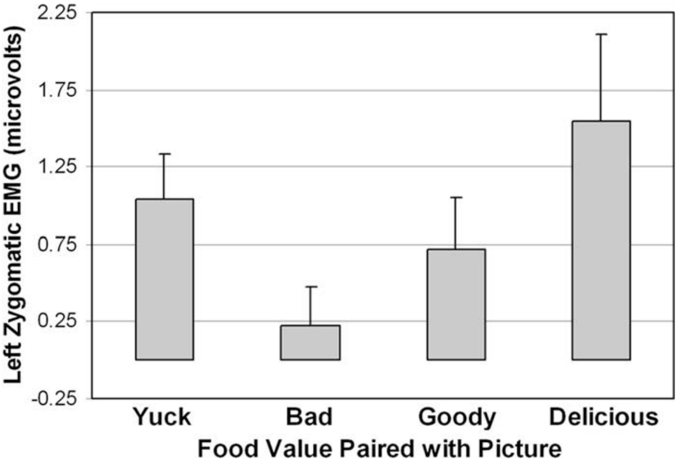

Fig. 7. Left zygomatic EMG responses to pictures paired with different food values for Experiment 2. Standard error bars represent 1 *SEM*.

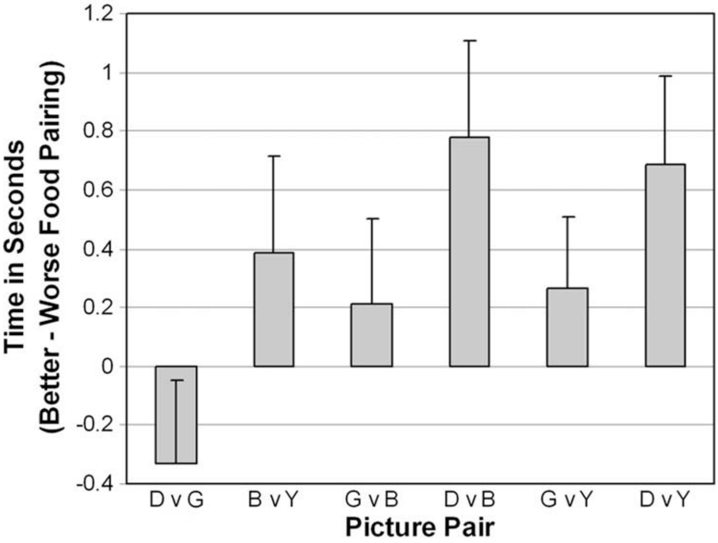

Fig. 8. Preferential looking for each picture pair for Experiment 2. Difference scores were computed by subtracting the amount of time spent looking at the picture paired with a worse taste from that paired with a better taste. Standard error bars represent 1 *SEM*.

pants in the present study exhibited the predicted linear relationship between picture pleasantness and corrugator response, and the quadratic relationship between picture pleasantness and zygomatic response (cheek muscles contract as stimuli become more pleasant, but also contract to very unpleasant stimuli). The facial muscle responses observed here correspond to those that were previously elicited by pictures holding intrinsic biological value or value acquired over several years (Fridlund et al., 1984; Lang et al., 1993; Schwartz et al., 1980). However, the current study shows that these expressions can adapt to reflect the current value of stimuli learned over a relatively small number of associative learning trials, and in a brief period of time.

Preferential looking was utilized as a convergent measure of picture preference. As predicted, participants spent more time looking at pictures associated with the food of higher value. Orienting towards or gazing at an object leads to its foveation for deeper sensory processing, and is important in establishing exposure to a stimulus and gathering information about its characteristics. In adults, work has shown that individuals look longer at items that they prefer (Shimojo et al., 2003; Karjbich et al., in preparation). More traditionally, preferential looking has been applied to test various aspects of visual discrimination abilities in monkey infants (Lee & Boothe, 1981; Mikami & Fujita, 1992; Regal, Boothe, & Teller, 1976), and to test vision and language abilities in human infants (Choi et al., 1999; Dobkins & Teller, 1996; Schafer & Plunkett, 1998; Southgate & Meints, 2000; Teller, 1979), under the premise that that infants look longer at something they can detect, or at something that is either novel or familiar. To our knowledge, this is the first use of preferential looking in a conditioning paradigm.

General discussion

Learning associated outcome

The present study demonstrates that human participants are able to learn the motivational value of different pictures according to the foods with which they are paired, and these values can vary in magnitude across the valence continuum. Although liking ratings for some of the close items (e.g., Goody and Bad in Experiment 2) were not statistically different according to standard cut-off criteria, the fact that pictures showed trends or statistical significance of the predicted relative values across two experiments and all five measures provides strong convergent evidence. These measures included preferential looking, corrugator muscle responses, zygomatic muscle responses, preference judgments, and liking ratings.

By measuring patterns of facial expression and eye movements to the pictures, the study indicates that human physiological responses also quickly adapt to reflect the different values of stimuli. This demonstration of learned physiological responses is an important step towards establishing a means of investigating corresponding behavioral and physiological measures of the current motivational value of a stimulus. Preferential looking has not previously been utilized in an adult human conditioning paradigm. The facial EMG data stand in contrast to the findings of previous associative learning experiments that have failed to demonstrate reliable facial EMG alterations to a stimulus following learning (Dimberg, 1987; Hamm et al., 1993; Sabatinelli et al., 1996; Schienle et al., 2001; Vaughan & Lanzetta, 1980). The current findings lend support to studies indicating that facial EMG does reflect learning of the negative motivational value of a stimulus (Flor et al., 2002; Bunce et al., 1999). The present study extends the findings to indicate that facial EMG can also reflect learning of positive motivational values of stimuli. Here it is of significance to note that the majority of studies that have been unsuccessful in demonstrating alterations in facial expressions as a function of learning have not used biologically relevant outcomes (Dimberg, 1987; Hamm et al., 1993; Sabatinelli et al., 1996; Schienle et al., 2001; Vaughan & Lanzetta, 1980). This could serve as an interesting point of investigation for future experiments.

The novel application of a logistic curve-fitting procedure in this study was particularly well suited for tracking learning of the different stimulus values. The procedure elucidated the dynamics of associative learning, while providing a measure of reliability with more traditional factorial analyses. The use of this procedure in conjunction with four different stimulus values enabled us to determine that the rate of acquisition of associations based on different food values is relatively invariant. Further investigation of this finding will be needed to demonstrate whether this is a common property of such associative learning experiments. It is possible that rate of acquisition is mediated by the degree of arousal elicited by each stimulus (as opposed to valence or value). This notion is supported by the relatively rapid acquisition of fear to previously neutral stimuli (LaBar, Gatenby, Gore, LeDoux, & Phelps, 1998). Despite the discrete and appropriate valence ordering of the stimuli utilized in the present experiment, it is possible that none of the stimuli elicited a high degree of arousal (relative to the ceiling of the human system). It is possible that a replication of the present experiment including parametric manipulations of hunger would demonstrate performance across different levels of arousal based on motivational state. Of some relevance to this issue is the fact that rats that are food-restricted do show different rates of learning for different outcome values (Chiba et al., 2002).

Substantial controversy exists over whether evaluative conditioning is an implicit or explicit process (De Houwer et al., 2001; Baeyens, De Houwer, & Eelen, 1994; Baeyens, Eelen, & Van den Bergh, 1990; Field, 2000; Lovibond and Shanks, 2002). For example, evidence for unaware evaluative conditioning has been obtained from careful studies that included between-subject control conditions to ensure that the results were based on associative learning and not other effects (Field & Moore, 2005; Hammerl & Fulcher, 2005). However, other work has convincingly found evaluative conditioning only after the moment at which participants could verbalize the stimulus-reward contingency (Purkis & Lipp, 2001). In the present experiment, there was a correlation between knowledge of picture–food contingencies and whether liking ratings for pictures were consistent with their outcome value. However, given that knowledge of picture–food contingencies were only investigated at the end of the experiment after extensive learning had occurred, it is not possible to determine whether participants' evaluations preceded or followed explicit awareness of the contingencies. Future experiments could track physiological responding during the learning process to determine whether there is evidence of conditioning prior to contingency awareness.

Regarding task demands, the consistency of findings across evaluative, physiological, and incidental measures suggest that stimulus acquisition of motivational values observed in the present experiments are based on actual hedonic value and not demand awareness. It is very unlikely that participants consistently consciously controlled their preferential looking behavior or their alterations in facial EMG, suggesting that our findings were unlikely due to participants complying with our hypotheses. Rather, the addition of physiological and looking measures to the standard evaluative measures, supports the existence of actual changes in the value of the initially meaningless pictures.

Conclusions

The current paper describes a useful framework for testing whether participants are able to associate outcomes with initially neutral stimuli. It suggests that, for different outcome values, stimulus-outcome pairings are learned equally well, and the latency and rate of this learning is consistent. Importantly, the experiment illustrates the feasibility of the application of measures of facial EMG and preferential looking to stimuli whose value has been acquired through association over short periods of time. Finally, this task holds great potential for examining the neural basis of affective associative learning and the central control of facial expressions, both potential components of "the smart gut".

Acknowledgments

Yoshio Takane supplied his multidimensional scaling program that was used in norming the visual stimuli used in this experiment. The authors thank Drs. Jan De Houwer, Laleh Quinn, and Mark G. Baxter for very helpful comments on this manuscript, and Dr. Leyla de Toledo-Morrell for her useful feedback on task design.

References

Ayers, E. D., & Powell, D. A. (2002). Multiple response measures following classical conditioning. *Journal of Neuroscience Methods, 114*, 33–38.

Baeyens, F., Crombez, G., Hendrickx, H., & Eelen, P. (1995). Parameters of human evaluative flavor conditioning. *Learning and Motivation, 26*, 141–160.

Baeyens, F., De Houwer, J., & Eelen, P. (1994). Awareness inflated, evaluative conditioning underestimated. *Behavioral and Brain Sciences, 17*, 396–397.

Baeyens, F., Eelen, P., & Van den Bergh, O. (1990). Contingency awareness in evaluative conditioning: A case for unaware affective-evaluative learning. *Cognition and Emotion, 4*, 3–18.

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*, 322–333.

Baeyens, F., Eelen, P., van den Bergh, O., & Crombez, G. (1992). The content of learning in human evaluative conditioning: Acquired valence is sensitive to US-revaluation. *Learning and Motivation, 23*, 200–224.

Berridge, K. C. (2000). Taste reactivity: Measuring hedonic impact in infants and animals. *Neuroscience and Biobehavioral Reviews, 24*, 173–198.

Bradley, M. M. (2000). Emotion and motivation. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (2nd ed., pp. 602–643). United States: Cambridge University Press.

Bunce, S. C., Bernat, E., Wong, P. S., & Shevrin, H. (1999). Further evidence for unconscious learning: Preliminary support for the conditioning of facial EMG to subliminal stimuli. *Journal of Psychiatric Research, 33*, 341–347.

Chiba, A. A., Quinn, L. K., Merzlyak, I. Y. (2002). Neural activity in the rat basolateral amygdala reflects the acquired motivational significance of visual objects. Program No. 284.8. 2002 Abstract Viewer/Itinerary Planner. Washington, DC: Society for Neuroscience.

Choi, S., McDonough, L., Bowerman, M., & Mandler, J. (1999). Early sensitivity to language-specific spatial categories in English and Korean. *Cognitive Development, 14*, 241–268.

Christenfeld, N., Glynn, L. M., & Gerin, W. (2000). On the reliable assessment of cardiovascular recovery: An application of curve-fitting techniques. *Psychophysiology, 37*, 543–550.

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*, 853–869.

Derryberry, D. (1989). Effects of goal-related motivational states on the orienting of spatial attention. *Acta Psychologia (Amst), 72*, 199–220.

Dimberg, U. (1987). Facial reactions, autonomic activity and experienced emotion: A three component model of emotional conditioning. *Biological Psychology, 24*, 105–122.

Dobkins, K., & Teller, D. (1996). Infant contrast detectors are selective for direction of motion. *Vision Research, 36*, 281–294.

Field, A. P. (2000). I like it, but I'm not sure why: Can evaluative conditioning occur without conscious awareness? *Consciousness and Cognition, 9*, 13–36.

Field, A. P., & Moore, A. C. (2005). Dissociating the effects of attention and contingency awareness on the evaluative conditioning effects in the visual paradigm. *Cognition and Emotion, 19*, 217–243.

Flor, H., Birbaumer, N., Hermann, C., Ziegler, S., & Patrick, C. J. (2002). Aversive Pavlovian conditioning in psychopaths: Peripheral and central correlates. *Psychophysiology, 39*, 505–518.

Fridlund, A. J., Schwartz, G. E., & Fowler, S. C. (1984). Pattern recognition of self-reported emotional state from multiple-site facial EMG activity during affective imagery. *Psychophysiology, 21*, 622–637.

Fulcher, E. P., & Cocks, R. P. (1997). Dissociative storage systems in human evaluative conditioning. *Behavioral Research Therapy, 35*, 1–10.

Gallagher, M., & Chiba, A. A. (1996). The amygdala and emotion. *Current Opinion in Neurobiology, 6*, 221–227.

Hamm, A. O., Greenwald, M. K., Bradley, M. M., & Lang, P. J. (1993). Emotional learning, hedonic change, and the startle probe. *Journal of Abnormal Psychology, 102*, 453–465.

Hammerl, M., & Fulcher, E. P. (2005). Reactance in affective-evaluative learning: Outside of conscious control? *Cognition and Emotion, 19*, 197–216.

Hatfield, T., Han, J. S., Conley, M., Gallagher, M., & Holland, P. (1996). Neurotoxic lesions of the basolateral, but not the central, amygdala interfere with Pavlovian second-order conditioning and reinforcer-devaluation effects. *Journal of Neuroscience, 16*, 5256–5265.

Holland, P. C., & Gallagher, M. (1999). Amygdala circuitry in attentional and representational processes. *Trends in Cognitive Science, 3*, 65–73.

Johnsrude, I. S., Owen, A. M., Zhao, W. V., & White, N. M. (1999). Conditioned preference in humans: A novel experimental approach. *Learning & Motivation, 30*, 250–264.

Karjbich, I., Armel, C., Rangel, A. (in preparation). Visual attention drives the construction and comparison of values in simple economic choice.

Kesner, R. P., & Williams, J. M. (1995). Memory for magnitude of reinforcement: Dissociation between the amygdala and hippocampus. *Neurobiology, Learning, & Memory, 64*, 237–244.

LaBar, K. S., Gatenby, J. C., Gore, J. C., LeDoux, J. E., & Phelps, E. A. (1998). Human amygdala activation during conditioned fear acquisition and extinction: A mixed-trial fMRI study. *Neuron, 20*, 937–945.

Lachnit, H., Lipp, O. V., & Gryschok, N. S. (2002). Probing the time course of nonlinear discriminations during human electrodermal conditioning. *Learning and Motivation, 33*, 269–283.

Lang, P. J. (1995). The emotion probe: Studies of motivation and attention. *American Psychologist, 50*, 372–385.

Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology, 30*, 261–273.

Lee, C., & Boothe, R. (1981). Visual acuity development in infant monkeys (*Macaca nemestrina*) having known gestational ages. *Vision Research, 21*, 805–809.

Levey, A. B., & Martin, I. (1975). Classical conditioning of human 'evaluative' responses. *Behaviour Research and Therapy, 13*, 221–226.

Lovibond, P. F., & Shanks, D. R. (2002). The role of awareness in Pavlovian conditioning: Empirical evidence and theoretical implications. *Journal of Experimental Psychology: Animal Behavior Processes, 28*, 3–26.

Mikami, A., & Fujita, K. (1992). Development of the ability to detect visual motion in infant macaque monkeys. *Developmental Psychobiology, 25*, 345–354.

Peinado-Manzano, A. (1989). Intervention of the lateral and central amygdala on the association of visual stimuli with different magnitudes of reinforcement. *Behavioral Brain Research, 32*, 289–295.

Pratt, W. E., & Mizumori, S. J. (1998). Characteristics of basolateral amygdala neuronal firing on a spatial memory task involving differential reward. *Behavioral Neuroscience, 112*, 554–570.

Purkis, H. M., & Lipp, O. V. (2001). Does affective learning exist in the absence of contingency awareness? *Learning and Motivation, 32*, 84–99.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II* (pp. 64–99). New York: Appleton-Century-Crofts.

Regal, D., Boothe, R., & Teller, D. (1976). Visual acuity and visual responsiveness in dark-reared monkeys (*Macaca nemestrina*). *Vision Research, 16*, 523–530.

Sabatinelli, D., Bradley, M. M., Cuthbert, B. N., & Lang, P. J. (1996). Wait and see: Aversion and activation in anticipation and perception [Abstract]. *Psychophysiology, 33*, 572.

Salinas, J. A., Packard, M. G., & McGaugh, J. L. (1993). Amygdala modulates memory for changes in reward magnitude: Reversible post-training inactivation with lidocaine attenuates the response to a reduction in reward. *Behavioral Brain Research, 59*, 153–159.

Schafer, G., & Plunkett, K. (1998). Rapid word learning by 15-month-olds under tightly controlled conditions. *Child Development, 69*, 309–320.

Schienle, A., Stark, R., & Vaitl, D. (2001). Evaluative conditioning: A possible explanation for the acquisition of disgust responses? *Learning & Motivation, 32*, 65–83.

Schwartz, G., Brown, S., & Ahern, G. (1980). Facial muscle patterning and participantive experience during affective imagery: Sex differences. *Psychophysiology, 17*, 75–82.

Sergent, J., & Takane, Y. (1987). Structures in two-choice reaction-time data. *Journal of Experimental Psychology: Human Perception and Performance, 13*, 300–315.

Shimojo, S., Simion, C., Shimojo, E., & Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nature Neuroscience, 6*, 1317–1322.

Southgate, V., & Meints, K. (2000). Typicality, naming, and category membership in young children. *Cognitive Linguistics, 11*, 5–16.

Takane, Y., & Sergent, J. (1983). Multidimensional scaling models for reaction times and same-different judgments. *Psychometrika, 48*, 393–422.

Teller, D. (1979). The forced-choice preferential looking procedure: A psychophysical technique for use with human infants. *Infant Behavior & Development, 2*, 135–153.

Todrank, J., Byrnes, D., Wrzesniewski, A., & Rozin, P. (1995). Odors can change preferences for people in photographs: A cross-modal evaluative study with olfactory USs and visual CSs. *Learning and Motivation, 26*, 116–140.

Vaughan, K. B., & Lanzetta, J. T. (1980). Vicarious instigation and conditioning of facial expressive and autonomic responses to a model's expressive display of pain. *Journal of Personality and Social Psychology, 38*, 909–923.