Chapter to appear in: J. M. Olson & M. P. Zanna (Eds.), *Advances in Experimental Social Psychology* (Vol. 43). San Diego: Academic Press.

Evaluative Conditioning: The "How" Question

Christopher R. Jones

The Ohio State University

Michael A. Olson

University of Tennessee

Russell H. Fazio

The Ohio State University

Please address correspondence to:

Russell H. Fazio
Department of Psychology
Ohio State University
1835 Neil Avenue
Columbus, OH 43210-1287
Phone: 614-688-5408
Fax: 614-688-5414
E-mail: fazio.11@osu.edu

Abstract

Evaluative conditioning (EC) refers to attitude formation or change toward an object due to that object's mere co-occurrence with another valenced object or objects. This chapter focuses on the "how" question, that is, the question of what cognitive processes intervene between mere co-occurrence and attitude formation or change. Though EC has typically been thought of as occurring through a single, albeit contentious, mechanism, we begin by pointing out that both the heterogeneity of EC methodologies and the abundance of inconsistent results suggest that multiple processes with different characteristics can produce EC. We describe how the earliest posited process of EC, Pavlovian conditioning or signal learning, is a valid mechanism of EC that appears to have operated in some experiments but is unlikely to have operated in others and also cannot account for various EC findings. We describe other mechanisms of EC, when they can be expected to occur, and what characteristics they have. We particularly focus our attention on a process model of EC we have recently introduced, the implicit misattribution model. Finally, we describe the implications of a multi-process view of EC, which we argue can help resolve theoretical controversies and further the application of EC as a practical intervention for influencing attitudes in various domains.

Evaluative Conditioning: The "How" Question

I. Introduction

Attitudes – our likes and dislikes – can be formed in a variety of ways. Sometimes we develop attitudes on the basis of an active and effortful process involving our consideration of the attributes that characterize an object, issue, or person. Indeed this attitude formation process lies at the heart of the expectancy-value framework (e.g., Fishbein, 1963). Obviously, we also develop attitudes on the basis of the outcomes that we experience upon interaction with the attitude object. Such interaction can produce a sense of satisfaction and pleasure, or a sense of dissatisfaction and even harm, and attitudes are shaped by these experiences (e.g., Fazio, Eiser, & Shook, 2004; Katz, 1960). We sometimes infer our attitudes from observation of our own behavior; our freely-chosen behavior provides diagnostic information regarding our likes and dislikes (Bem, 1972). It is also well-established that attitudes can be developed through conditioning as a result of the mere co-occurrence of an attitude object with other liked or disliked stimuli. Such evaluative conditioning forms the focus of this chapter, with particular attention to what we refer to as the "How" question: By what mechanisms does evaluative conditioning arise?

Evaluative conditioning (EC) refers to attitude formation (or change) due to an object's pairing with positively or negatively valenced stimuli. Adopting the terminology of Pavlovian conditioning, the former is referred to as the conditioned stimulus (CS) and the latter as the unconditioned stimuli (US). Thus, evaluative conditioning (EC) involves transfer of the valence associated with the US to the CS. But, just how does this "transfer" happen? The field has witnessed a rather remarkable resurgence of interest in evaluative conditioning over the last ten years or so. Many articles employing some variation of a conditioning procedure have appeared,

as well as influential review papers (e.g., De Houwer, Thomas, & Baeyens, 2001) and special issues of journals (*Cognition & Emotion*, 2005). Yet, many mysteries continue to surround evaluative conditioning and, in our view, none is more central than the question of underlying mechanism.

A. Two classic experiments

To introduce our consideration of the process or processes by which evaluative conditioning might arise, let's briefly review two classic experiments. Almost every historical summary of the literature concerning the formation of attitudes through conditioning includes presentation of Staats and Staats (1958) and Zanna, Kiesler, and Pilkonis (1970). The Staats and Staats work is considered one of the very first demonstrations of such attitude formation. The procedure involved a clever cover story in which participants were led to believe that the experiment concerned the simultaneous learning of words presented via different modalities. They first were asked to learn a list of words that were presented visually, and were tested on their learning of the list. They then did the same with a list of words presented orally, all of which set the stage for the critical phase of the experiment which was portrayed as an assessment of participants' ability to learn via both visual and auditory channels at once. During this phase, participants were exposed visually to a set of nationality names, including the CS nationalities *Dutch* and *Swedish*. Approximately one second after the nationality appeared on the screen, the experimenter announced a word aloud. Most of these latter words, none of which were repeated, were evaluatively neutral, connoting little valence (e.g., *chair, with, twelve*). Included, however, were a few positive words (e.g., *gift, sacred, happy*) and a few negative words (e.g., *bitter, ugly, failure*). These US words were systematically paired with the two CS nationalities such that one always appeared with positive words (the CSpos) and the other with negative words (the CSneg).

Thus, the conditioning trials were embedded within a stream of visually presented nationality names and orally presented words.

When the conditioning phase was completed, the participants were first asked to recall the words that had been presented visually and then to evaluate them, presumably because how they felt about those words might have affected their learning. The conditioning was successful. The nationality that had been paired with the more positive US was rated as more pleasant than the one paired with the negative US.

The Staats and Staats (1958) experiment proved controversial, largely due to the possible role of experimenter demand characteristics. Some (e.g., Page, 1974) argued that participants may have become aware of the CS-US pairings, inferred that the experimenter expected their ratings to vary as a function of the pairings, and then simply complied with the presumed "demand" of the situation. It was this very issue that motivated the experiment by Zanna et al. (1970), who made use of the now common two-experiment ploy to separate collection of the dependent measure from the conditioning phase of the experiment. The ratings of the CS were obtained by a second experimenter in the context of an ostensibly different experiment, thus alleviating concerns about demand characteristics.

However, as shall be evident soon, there is another reason we wish to highlight the Zanna et al. work and, in particular, the conditioning procedure that was employed. The procedure began by leading participants to believe that the experiment aimed to test the effectiveness of various physiological measures as indicators of arousal. Electric shock of varying duration was to be delivered for the purpose of examining whether the physiological measures that were being recorded were sufficiently sensitive to detect the precise moments of shock. Although the duration of the shock on any given trial was to be randomly determined, participants were told

that the onset of the shock would be signaled by a particular word and the end of the shock period would be signaled by yet another word. Presumably, these signal words would allow the participants to relax between the shock trials and, hence, have their physiological reactions return to baseline. Participants were explicitly told what the signal words were at the beginning of the procedure. The signal words themselves were embedded within a series of words that were played on an audio-recorder, with the number of filler words varying from trial-to-trial. Thus, while listening to a stream of words, participants knew that one particular word (e.g., light) would signal shock onset whereas another (e.g., dark) would signal shock offset. After 25 such trials, participants were told that the experiment required a 15-minute rest period. It was during this time that they were introduced to a second experimenter who had them complete, among other tasks, a word-meaning survey. Included in the list of words being evaluated were the signal words (light/dark). Relative to a control condition, the word that signaled shock onset shifted in a less favorable direction, whereas the word that signaled offset assumed a more favorable meaning.

As should be obvious, and as we wish to emphasize, the procedures employed by Staats and Staats (1958) and by Zanna et al. (1970) differ markedly. Yet, they do share one basic commonality. In both, the participant was exposed to a systematic pairing of CS and US; the CS and US were presented in a manner that can be characterized as involving spatio-temporal contiguity. Otherwise, they bear little resemblance to one another. First, the two procedures differ with respect to whether anything hedonically consequential is happening to the participant. The Zanna et al. procedure involved the experience of electric shock, whereas little of any hedonic consequence occurred in the Staats and Staats procedure. Second, the temporal relationship between the CS and US differs. In the Zanna et al. procedure, a spoken CS word

immediately precedes the onset of the US. In the Staats and Staats procedure, the CS and US presentations overlap. The visual CS appear before the US, but after the first second of a 5 s CS presentation, the US is spoken by the experimenter. Third, the number of unique US differs. The Zanna et al. procedure involved a single hedonically consequential event as the US -- the experience of an electric shock that was repeatedly paired with the CS. In the Staats and Staats procedure, none of the orally-presented US words was repeated; a variety of US words appeared with each CS. Only the valence of the US systematically covaried with the presentation of the visually-presented CS nationality names. Fourth, the apparent relationship between the CS and US differs greatly. In the Zanna et al. procedure, the shock's onset and offset were signaled by CS whose very relation to the shock was explicitly communicated. In contrast, the Staats and Staats procedure involved a CS and US from separate, presumably unrelated tasks; the procedure was designed specifically to obscure the CS-US relationship. Both procedures yielded evidence of evaluative conditioning; the CS came to be evaluated differentially as a function of the experience. However the question we wish to raise is whether the two procedures achieved this outcome via the same mechanism. Just how reasonable is it to view the underlying process to have been the same in these two classic studies?

B. The "How Question"

The aim of this chapter is to further understanding of this basic question regarding the processes that underlie evaluative conditioning. The question is not only of fundamental importance, but in our view it also relates centrally to a number of longstanding issues regarding EC - two of which are useful to highlight now. One controversy that has long characterized the EC literature concerns the role of contingency awareness, i.e., the extent to which individuals are aware of the systematic nature of the CS-US pairings. Is such awareness necessary for EC to

occur? The literature could not be more muddled about this matter (for reviews, see De Houwer, Thomas, & Baeyens, 2001; Field, 2000a; Shanks & Lovibond, 1994). There are data to suggest that EC does not occur without such awareness (e.g., Pleyers, Corneille, Luminet, & Yzerbyt, 2007; Pleyers, Corneille, Yzerbyt, & Luminet, 2009; Purkis & Lipp, 2001), that contingency awareness is unrelated to the extent to which EC is observed (e.g., Baeyens, Eelen, & Van den Bergh, 1990; Baeyens, Hermans, & Eelen, 1993), and even that contingency awareness may interfere with EC (e.g., Fulcher & Hammerl, 2001; Walther & Nagengast, 2006).

A second matter that has received considerable attention is the apparent fragility of EC. Our own experiences in our laboratories certainly would lead us to characterize EC as a seemingly ephemeral effect. Indeed, that appears to be a common observation among the community of scientists who have engaged in EC research. As De Houwer, Baeyens, and Field (2005) noted, there are many reports, both published and unpublished, of failures to observe EC, leading them to conclude: "EC is certainly not a robust research finding" (p.163). Indeed, Rozin, Wrezesniewski, and Byrnes (1999) discussed this very issue in an article aptly titled "The Elusiveness of Evaluative Conditioning."

In our view, these two issues simply point to the fact that the boundary conditions that characterize EC and the underlying mechanism(s) by which it is produced are poorly understood at this point in time. Clearer understanding of the potentially multiple processes that may be involved would serve to identify conditions that are optimal for the development of EC, and contingency awareness may play different roles in different processes. It is for this reason that we, like others (e.g., De Houwer, 2007), believe it to be advantageous to define EC, not as a single process, but as an *effect*. EC is an outcome -- attitude formation (or change) due to an

object's (CS) spatio-temporal contiguity with positively or negatively valenced stimuli (US).[1] This transfer of valence may occur through multiple mechanisms, just as suggested by it having been produced by such vastly different procedures in the two classic experiments we considered earlier.

1. A partial answer: Pavlovian conditioning

It may seem odd for us to have asserted that the mechanism or mechanisms underlying EC are poorly understood given the decades of theory and research devoted to human and animal conditioning. This literature is certainly relevant to the matter of EC and provides a clear framework regarding one potential mechanism by which EC might occur. Contemporary perspectives portray Pavlovian conditioning as a form of signal or expectancy learning in which the CS signals the pending occurrence of the US (Rescorla, 1988). As such, Pavlovian conditioning depends not only on the number of pairings of the CS and US, but on the statistical regularity of their co-occurrences. US occurrence in the absence of the CS interferes with learning, and CS occurrence in the absence of the US promotes extinction. Obviously, the statistical covariation is the key to a stimulus effectively signaling that a hedonically consequential event is about to occur. Moreover, any contingency awareness can only enhance the utility of this signal value.

At this point, it must be apparent that the Pavlovian conditioning model corresponds well to the procedure employed by Zanna et al. (1970). Knowledge developed that a specific word signaled the pending electric shock and this signal value was regularly validated throughout the

[1] The only element of the definition that includes any restrictions regarding process is the matter of spatio-temporal contiguity. To be considered EC, the attitudinal outcome needs to stem from that CS-US contiguity. Otherwise, any attitude formation or change process, including persuasion and the deliberative analysis of the likely outcomes produced by interaction with an attitude object, would fall under the rubric of EC. So as not to render the term meaningless, we regard EC as stemming from the mere spatio-temporal contiguity of the CS and US.

procedure. Later experiments have employed similar procedures and also have successfully demonstrated attitude formation through such signal learning (e.g., Cacioppo, Marshall-Goodell, Tassinary, & Petty, 1992). Signal stimuli that are associated with a consequential event that is negative in nature come to be viewed less favorably, whereas those signals that predict a pleasant experience are viewed more positively.

2. Can Pavlovian conditioning explain all the data?

Yet, this mechanism does not appear to characterize the Staats and Staats procedure. In contrast to receiving an electric shock, the oral presentation of such words as *gift* and *failure* hardly seems a hedonically consequential event. How plausible is it to maintain that simply hearing such words, amidst a lengthy list of unrelated ones, would be consequential enough for signaling stimuli to be sought, noticed, or in any way appreciated by an information-processing system?

With that question in mind, we ask that the reader remain open to the possibility that EC may be produced by multiple mechanisms, only one of which is the signal learning process delineated by the Pavlovian conditioning perspective. This possibility, and its implications, will constitute major sections of this chapter. However, before turning to the mechanism that might underlie EC in the Staats and Staats (1958) procedure, it will be useful to review a series of experiments that we recently conducted concerning EC. The experiments all involve a paradigm that we refer to as "video surveillance" (Olson & Fazio, 2001). Although we did not recognize it when the research was initiated, with the benefit of hindsight, our consideration of the process question has led us to view the paradigm as a modern equivalent of the Staats and Staats procedure. First, it shares with the paradigm of Staats and Staats that great care is taken to obscure the fact that it is intended to influence attitudes. Second, it involves the overlapping

presentation of a given CS with multiple distinct US, though the surveillance procedure presents both CS and US in one modality (visual) whereas the Staats and Staats procedure is cross-modal (visual CS and auditory US).

II. The Surveillance Procedure

A. Overview

In developing this EC paradigm, our goal was to produce a change in respondents' attitudes toward an object through its repeated pairings with other valenced stimuli while minimizing demand and contingency awareness. In other words, we wanted to produce an EC effect and rule out alternative explanations based on higher-order processes. We were not so interested in elucidating a specific mechanism that we thought might account for it, as the inconsistent findings and lack of theorizing in the literature seemed to suggest that such a search was premature at the time. However, we looked to the vast annals of classical conditioning research for inspiration, and had many conversations with our colleagues trained in the tradition of learning theory (many of whom employed animal models to investigate signal learning in the tradition of Rescorla-Wagner) as we tried to arrive at a paradigm that might produce reliable conditioning implicitly. As we progressed, we wondered whether we were indeed in search of a different animal than that studied by the traditional learning theorists. For instance, a large body of literature overwhelmingly implicated forward conditioning (where the CS precedes the US) as a more effective means of generating conditioned responses than either backward conditioning (where the US precedes the CS) or simultaneous conditioning (where the CS and US are presented together; Malone, 2002). Our intuitions suggested otherwise; although forward conditioning made sense to us from the perspective of an organism oriented toward predicting consequential events in its environment (i.e., signal learning), as attitudes researchers we were

more interested in the development of evaluative responses than anticipatory responses. Our sense was that the affect from the valenced stimulus (the US) somehow transferred or "bled" onto the attitude object (the CS). Armed with such imprecise metaphors, we decided that a simultaneous presentation of CS-US pairs would provide the greatest likelihood of such transference.

B. General method

Thus, the particular characteristics of the surveillance procedure were chosen with aims including to minimize contingency awareness but not on the basis of characteristics known to facilitate PC. The surveillance procedure begins with instructions to participants. They are asked to imagine that they are playing the role of a security guard monitoring for unusual activity (or in other words, to imagine engaging in surveillance). Participants are informed that they will view a series of images on a computer screen, and their job is to be vigilant and to respond as quickly as possible whenever a target item appears by pressing the space bar as quickly as possible. Thus, the experience of the participant is akin to "video surveillance" in which events occurring elsewhere are viewed on a screen. Participants are warned that these targets will appear in the context of many "distracter" stimuli intended to make the task more difficult and simulate the complex nature of a real environment under surveillance. A target packet is given to participants containing the cartoon characters that serve as target stimuli. Each cartoon, both on the packet and during the procedure, includes a name label at the bottom of the picture.

When the task itself begins, participants are presented with a stream of visual stimuli. Each trial, usually around 1.5 s in length, involves the presentation of zero to two stimuli, most of which are "filler" trials involving relatively neutral words (e.g., "concrete," "smell," or "Poliwag" the name of a filler cartoon character) and images (an airplane, a basket, Poliwag the

cartoon). These filler trials are randomly ordered and include a number of "blank" trials (no stimuli) that give the presentation a non-rhythmic quality and make it difficult to determine exactly when stimuli will appear. Embedded in this stream of filler trials are two other types of trials: target and conditioning trials. Target trials involve the presentation of a target cartoon in word form (name label only) or image form (picture and name label). Sometimes a filler stimulus accompanies the target (i.e., is presented at the same time). It is these target trials on which participants are supposed be vigilant for and respond with a key press. Target trials are frequent, appearing every 10 seconds or so; by the end of the procedure participants will have had the opportunity to respond to 50 targets. Conditioning trials, embedded at fixed intervals, involve the presentation of a CS cartoon (with name label) and a valenced US word (e.g., "fantastic," "awful," or "awesome") or image (e.g. puppies, a cockroach, or a picturesque mountain). Two CS are used, and these particular cartoons were those identified in pretesting as being most neutral with the least variability in evaluative ratings. One CS (the CSpos) appears always and only with positive stimuli and the other (the CSneg) appears always and only with negative stimuli. Which CS cartoon appears with US of a given valence is counterbalanced across participants through random assignment, ensuring that any post-procedure differences in evaluation of the CSpos and CSneg is due to conditioning rather than differences inherent to the two CS cartoon characters. The US stimuli are also selected on the basis of pretesting confirming that they are unambiguously positive or negative.

Attitude theorists generally agree that attitude creation is easier than attitude change; hence, objects novel to most undergraduate research participants were chosen as CS. We also suspected that our efforts to influence peoples' gut-level affective reactions would be more successful if the attitude objects seemed to warrant gut-level responses (Croizet & Fiske, 2000).

In contrast to most stimuli employed in implicit learning paradigms in the cognitive tradition (e.g., Ashby & Maddox, 2005), these cartoon characters have anthropomorphic qualities, and the colleagues to whom we vetted our ideas indicated that they could envision themselves liking or disliking them. We had little theory or prior research to guide our selection of US. The question of whether to use images or words confronted us early on, and with little reason to rule out one or the other, we chose both. We looked to the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 2008) for positive and negative images, and to our history with evaluative adjectives (Fazio, Sanbonmatsu, Powell, & Kardes, 1986) for words.

The surveillance procedure entails five blocks of trials. Each block has a different target stimulus (or, more recently two stimuli, which makes it somewhat more difficult for participants to use a simple cue like color to identify a target and is intended to ensure non-target objects are processed). Each block involves 86 trials, including 10 target trials and 8 conditioning trials (4 CSpos and 4 CSneg). Thus, each CS appears with 20 different US of a given valence across the five blocks. After each block, participants have the opportunity to rest briefly and to familiarize themselves with the next target(s). Immediately after the procedure, participants' liking for the two CS is assessed followed by an assessment of contingency and demand awareness.

Various aspects of this procedure are intended to minimize the likelihood that participants will notice with what stimuli the CS appeared (contingency awareness) or realize that these pairings are intended to influence attitudes (demand awareness). The cover story and surveillance task require constant attention to the display, ensuring that the CS-US pairings are perceived, but they do so in a way that draws no particular attention to these critical trials and justifies their appearance. The participant's task is to be vigilant for the target's appearance, which has nothing to do with the matter of two stimuli appearing on the screen at the same time.

In addition, the usage of multiple, non-repeating US greatly reduces the likelihood that participants will form associations in memory between a CS and any particular US, which would require one-trial learning.

C. Research findings

The most basic finding from the surveillance procedure is that it is effective in producing attitudes toward the CS. First, this has been demonstrated using explicit measures of attitudes. In our first published work using the procedure (Olson & Fazio, 2001), participants underwent the procedure as described above, followed by a measure of their attitudes toward the CS and a number of fillers using a Likert-type scale. Because we were concerned that participants would not report any conditioned attitudes for lack of a "reason" for feeling any particular way about them (Croizet & Fiske, 2000), they were encouraged to "go with their guts" and not think much before responding to each item. To compute an EC effect, participants' responses to the CS paired with negative items was subtracted from their ratings of the CS paired with positive items. Results indicated a positive and reliable mean conditioning effect. Alternatively, in a number of studies liking has been assessed through a forced-choice procedure in which participants are presented a number of trials on which two cartoon characters appear, and they must select the one that they prefer (Jones, Fazio, & Olson, 2009). This method has the advantage of promoting the fast expression of "gut" preferences and it too has shown that participants come to like the CSpos more than the CSneg.

These conditioning effects have also been demonstrated using implicit measures of attitudes. Such implicit measures are believed to tap evaluative processes that are relatively more automatic than direct, explicit measures (see Fazio & Olson, 2003; De Houwer, Teige-Mogicemba, Spruyt, & Moors, 2009; Olson & Fazio, 2009), most notably in the sense of their

being relatively difficult to control and/or assessing spontaneous attitude activation. Thus such measures provide stronger evidence that observed EC truly reflects changes in the attitudes associated with the CS in memory. Olson and Fazio (2001, Experiment 2) found that the surveillance procedure produced an EC effect simultaneously on both an explicit rating measure and on an implicit association test (IAT) designed to tap any newly-formed associations to the CS that formed: participants were quicker to map the CS paired with positive items and "Pleasant" onto the same response key (and the CS paired with negative items and "Unpleasant" onto another) than the reverse. In another experiment, the CSpos and CSneg served, following conditioning, as primes in a subliminal evaluative priming task (Olson & Fazio, 2002). Even though the CSpos and CSneg were presented briefly and masked during attitude measurement so that participants could not identify what they were seeing, they influenced how easily participants could identify the valence of adjective targets. Participants more quickly identified positive adjectives and more slowly identified negative adjectives following the CSpos prime compared to CSneg prime, indicating that these stimuli had acquired attitudes that were capable of automatic activation upon their perception even when that perception was below the threshold of consciousness.

Evidence of EC on evaluative priming measures also has also been observed using stimuli towards which participants could be expected to already possess attitudes (socially significant attitudes, in fact). Typically, a substantial proportion of participants show relatively negative attitudes toward Black people on evaluative priming measures (e.g. Fazio, Jackson, Sanbonmatsu, & Kardes, 1995). In Olson and Fazio (2006), images of Black and White people in various occupational settings replaced the cartoon CS and targets; otherwise, the procedure was essentially identical. Mixed in among the filler images, experimental participants were

exposed to systematic pairings of a number of Black faces with positive stimuli and a number of White faces with negative stimuli. Control participants were exposed to the same stimuli singly, i.e., without any systematic pairings. Participants then completed a priming measure of automatically-activated racial attitudes, a measure with a history of proven predictive validity (see Olson & Fazio, 2009, for a review). Results on the measure confirmed that positivity was more likely to be evoked following Black relative to White faces among experimental participants; control participants evidenced the reverse pattern more typically seen in this work. This conditioning effect was observed both immediately and, in another experiment, following the passage of a two-day interval between the conditioning phase and the administration of the evaluative priming measure. This latter finding is particularly significant because it demonstrates that the surveillance procedure can produce attitudes that are stable enough to persist beyond the initial laboratory session.

In all of the experiments just described, EC effects were demonstrated in the absence of contingency awareness. In almost all cases, a funneled post-experimental questionnaire probed participants' ability to report what valence of stimuli had appeared with the CS. This measure begins rather vaguely, probing participants' general suspicion about the procedure and culminates by specifically asking participants to report whether they had noticed anything about the stimuli that appeared with the CSpos and CSneg. Though some participants were able to correctly report that at least one CS appeared with US of a given valence (this proportion ranged across samples from 2% to 18%, excluding a condition of Jones, Fazio, & Olson, 2009, Experiment 5 in which particularly strong US stimuli were used and produced a high degree of contingency awareness), these EC effects were still evident after these participants had been excluded from analyses.

Other experiments employed (Olson & Fazio, 2001, Experiment 1) (or more frequently *also* employed) a recognition measure of contingency awareness. Participants reported their degree of certainty that a CS had appeared with various stimuli, including several stimuli with which the CS had indeed appeared and several stimuli opposite in valence with which it had never appeared. Participants were no more certain of the former true instances than the latter foils. In other words, performance on this recognition measure was no better than chance, which is very understandable in light of the participants' task demanding not attention to the simultaneous appearance of any two stimuli, but vigilance for the appearance of a specific target. Further, participants' accuracy on this task did not predict their ratings of the CS. In other words, this explicit recognition measure corroborates evidence from the post-experimental questionnaire[2] that the surveillance procedure creates attitudes that are not rooted in participants' observation that a CS appeared with a particular US or type of US. We usually exclude contingency-aware participants from subsequent analyses to insure that our focus is on implicit associative learning, although their exclusion typically does not alter the pattern of results.

III. The Implicit Misattribution Model

A. Relation to the Pavlovian conditioning account

Given these findings, we return now to the "how" question. What sort of mechanism is consistent with the method we have described and the characteristics of EC as it manifests through the surveillance procedure? It's not, it would appear, an attitudinal byproduct of classical Pavlovian conditioning (PC), the phenomenon that has guided much thinking about and

[2] In addition to supporting the general conclusion that EC can occur without contingency awareness, when included in the same experiment, responses to the post-experimental questionnaire and the recognition memory test generally correspond in terms of their identification of participants who are aware of CS-US contingencies. In multiple experiments, those categorized as unaware on the basis of the post experimental questionnaire performed at chance levels on the recognition test. However, those classified as aware showed better recognition memory than those considered unaware. The correspondence between these measures provides further evidence of their validity.

the terminology of EC. PC involves learning that an event in a given context signals the arrival of another hedonically relevant event. It crucially involves the awareness of the relation between the two objects; some have argued that it cannot take place without it. Data from the surveillance procedure are consistently at odds with this characteristic of PC.

Other researchers have also demonstrated that EC can look very much unlike PC. The ease with which learning through PC occurs is a function of the degree of consistency with which the CS predicts the US. A CS is of little use for predicting a US if the two do not reliably co-occur. Baeyens, Hermans, and Eelen (1993), however, found that the degree of CS-US co-occurrence did not determine the strength of EC. Adding 10 CS-only trials to 10 CS-US pairings did not reduce EC compared to 10 CS-US pairings (perfect covariance). Instead, the absolute number of pairings tends to predict the magnitude of EC (see De Houwer, Thomas, & Baeyens, 2001 for a review). Similarly, PC is subject to the well-known phenomenon of extinction. After acquisition of PC, presenting the CS alone gradually extinguishes the effect as the CS loses signal value. Several studies have shown EC to be rather robust in this sense, diminishing little or not at all following extinction trials (see De Houwer, Thomas, & Baeyens, 2001). Others, however, have found that EC can be sensitive to extinction (e.g., Lipp, Oughton, & LeLeivre, 2003), begging the question of whether the same process is mediating the effect of spatio-temporal co-occurrence on attitude change.

B. A novel mechanism: Implicit misattribution

1. Misattribution in social psychology

If not signal learning, then what can account for EC such results? In an effort to explain how EC occurs in the surveillance procedure and likely some other methods, we have proposed a novel process model of EC, the Implicit Misattribution Model (IMM). According to the IMM, a

response evoked by the US is misattributed, without awareness, to the CS, lastingly influencing an attitude toward the CS. In other words, a confusion of the source of a subjective experience occurs. We do not mean to suggest that a *feeling of confusion* occurs; rather, the *act of confusing* which of multiple co-occurring objects in one's environment is evoking a response is sufficient to produce attitude change. Our model draws on theory and research that identifies attributional processes as important to the experience of affect. Indeed, some emotion theorists have argued that what we experience as emotions are the product of the fundamental cognitive and physiological processes of valence and arousal being given fuller psychological significance by the attribution of these basic, core feelings to particular causes (Russell, 2003; Schachter & Singer, 1962). The specific causes of our feelings are, however, not always obvious.

Extensive research in the domain of attribution has in fact repeatedly demonstrated misattribution of affect. In one well-known example, Schwarz and Clore (1983) phoned participants in their homes and asked them how satisfied they were with their lives. Participants reported greater general life satisfaction on pleasant, sunny days than unpleasant, rainy ones. Apparently, participants misattributed their transient moods caused by the weather to more chronic general life satisfaction. When participants in another condition were reminded of the weather just before the life satisfaction question, however, no influence of weather was found. When that cause of one's current state was made salient, it was correctly identified as a source. Many "classic" studies illustrate the fundamental point that individuals are prone to misattribute various feelings to salient and seemingly appropriate objects in their environments (e.g. Dutton & Aron, 1972, Schacter & Singer, 1962; Storms & Nisbett, 1970; Zanna & Cooper, 1974; Zillman, Katcher, & Milavsky, 1972).

More recently, interesting examples of affective misattribution have been observed as a consequence of the experience of *value from regulatory fit* (Higgins, 2005). This refers to the notion that congruency between the type of goal one has and the means of pursuing it produce positive feelings. For example, a *prevention* goal focused on the avoidance of an undesirable end-state would fit better with goal pursuit entailing *vigilance against* particular occurrences than e*ager pursuit for* other occurrences. These positive feelings produced by regulatory fit, or pursuing a goal "the right way," can be misattributed. Cesario, Grant, and Higgins (2004) found that framing message arguments in a manner congruent with recipients' regulatory state increased the persuasive impact of the message, arguing that the positive feelings produced by fit were misattributed to the arguments themselves. In support of this account, they demonstrated that such effects of fit on persuasion could be eliminated by drawing attention to the actual source of those positive feelings. This feeling of rightness from fit can also be misattributed to one's progress in goal pursuit, influencing individuals' "stop rules" or decisions to cease goal pursuit on the basis of sufficient progress (Vaughn, Malik, Schwartz, Petkova, & Trudeau, 2006). Supporting the role of attribution, the effect was eliminated by drawing attention to an earlier event as a source of positive feelings derived from regulatory fit.

In a similar fashion, the experience of perceptual fluency has been shown to influence aesthetic preferences (Reber, Schwarz, & Winkielman, 2004). Processing fluency, which has been argued to generally indicate familiarity and safety, appears to inherently elicit positive affect (Winkielman & Cacioppo, 2001). This fluency can be mistaken (e.g., when the fluency is not due to previous experience) as indicative of the object's value. In one experiment (Reber, Winkielman, & Schwarz, 1998), stimuli were preceded by matching or non-matching contour primes, easing or complicating processing of the target stimuli, respectively. Those preceded by

a matching prime were subsequently deemed more aesthetically pleasing. Apparently, the affect elicited by the experience of processing a stimulus fluently can be misattributed to the object itself. Such research is particularly notable with regard to present concerns, because the misattribution of affect from fluency was demonstrated to influence object evaluation per se.

The IMM, in brief, states that one way that the co-occurrence of two objects can lead to attitude change is through the misattribution of an attitude evoked by the US to the CS. Before discussing the model in more detail, we will review other research demonstrating that misattribution can influence evaluation. Murphy and Zajonc (1993) first demonstrated that subliminally presented valenced images influence how positively individuals evaluate subsequently presented neutral symbols. Indeed, this basic effect can be utilized to produce an implicit measure of attitudes by quantifying the extent to which objects of interest influence subsequent evaluative judgments of neutral objects (Payne, Cheng, Govorun, & Stewart, 2005). Payne and colleagues, in particular, have described this effect of primes on subsequent evaluative judgments in terms of misattribution. These studies, however, only demonstrated effects of primes on immediate judgments. If such misattributions can also lastingly influence evaluations, it would seem that this process suffices to produce evaluative conditioning.

Thus, the IMM was at the outset plausible insofar as misattribution of affect has been a widely observed phenomenon, and previous research suggests that such misattributions can at the very least influence evaluations transiently. In addition, research on automaticity rendered plausible the notion that this process can occur implicitly, that is without individuals' awareness—not just without awareness that a *mistaken* attribution is being made (indeed, the idea of *knowing* oneself is *mis*attributing an event to a cause is paradoxical) but without awareness that attribution is occurring at all. First, we will note the phenomenon of automatic

attitude activation (Fazio, Sanbonmatsu, Powell, & Kardes, 1986). Upon encountering an unconditioned stimulus, individuals may experience automatic activation from memory of an attitude towards that object. The likelihood of this occurrence is partially a function of the associative strength in memory between the object and the attitude. Automatic attitude activation is known to be efficient and spontaneous, that is, quick and not dependent on an explicit goal of assessment. Similarly, individuals constantly attribute events to particular causes without intending to make such inferences (Weiner, 1985) or being aware of them. For example, the effect of perceptual fluency on liking appears to operate without conscious awareness (Reber, Schwarz, & Winkielman, 2004). Not only might individuals not intend to make attributions, they may be unaware of doing so. Considerable evidence exists to support the contention that many mental processes operate outside of conscious awareness, attribution among them, and while the *product* of these processes (a changed attitude in this case) may be available consciously, the operations that produced it are not (Nisbett & Wilson, 1977).

In sum, implicit misattribution was plausible as a source of EC because misattribution is common and the processes of attitude activation and attribution are known to function with a high degree of automaticity. Therefore, one need not experience uncertainty about the source of subjective experience and engage in a thoughtful but mistaken search of one's environment for a plausible cause to produce EC. Instead, we suspected that such subjective experience would be linked to external causes largely automatically and at an early stage of perceptual-cognitive processing. In other words, the myriad attitudes that are activated automatically in daily life can become fodder for attribution processes that also occur automatically, creating the possibility of transfer of affect from a known attitude object (US) to a contiguous object (CS), all in the absence of awareness of the process' operation.

2. Feature migration: An analogue

The proposed process of implicit misattribution does, in fact, have an interesting analogue in the science of perception called feature migration or illusory conjunction (Treisman & Schmidt, 1982). The essence of the phenomenon stems from the fact that the process of object identification requires the integration of more fine-grained features, and this assembly process can sometimes err. Feature migration occurs when a feature or element of an object is misperceived as belonging to a different object. For example, Treisman and Schmidt (1982) recounted an anecdote in which a man thought he recognized a colleague across a busy street. Soon, however, "he realized that the black beard belonged to one passerby and the bald head and spectacles to another" (p. 109). They confirmed in laboratory experiments that such feature migrations occur readily under certain conditions. For example, a particular configuration of a red *T*, a green *X*, and a blue *N* often led participants to report having seen a blue *T*. Feature migration was predicted on the basis of the feature-integration theory of visual attention (Treisman & Gelade, 1980), which posits that primitive features (e.g., color, orientation) can be processed in parallel and pre-attentively. On the other hand, the binding of features into perceptions of distinct objects occurs serially and requires directed attention. This resource dependent process is more error prone than feature identification. Notably, feature migration is not limited to visual features; it has been observed, for example, in the auditory domain between pitch and duration (Thompson, Hall, & Pressing, 2001).

We do not wish to argue that feature migration and implicit misattribution are one and the same. An attitude is not the same as a primitive perceptual feature. The parallels between feature migration and implicit misattribution, however, are interesting. Both depend on a high degree of spatio-temporal overlap between objects. Also, evaluative responses to objects occur

very quickly, at relatively early stages of perceptual-cognitive processing. Zajonc (1980) famously argued for affective primacy, the idea that evaluative appraisals precede the activation of other knowledge. Evidence that attitudes, like features, are activated pre-attentively comes from research by Roskos-Ewoldsen and Fazio (1992) who demonstrated that accessible attitudes automatically direct visual attention to hedonically relevant objects presented peripheral to one's visual focus. This occurred even when such guidance of attention towards these objects was detrimental to participants' task performance. In other words, attitude-evoking objects automatically attract attention when they enter the visual field, more so than objects that are not evaluatively significant. In sum, it appears that there is reason to believe that attitude activation, like feature identification, precedes perceptual binding, suggesting that when this process of object delineation errs, attitudes, like features, may be confused. The attitude may be "assembled" with an object that did not evoke it.

Evidence from neuroscientific research converges with this conclusion. The iterative reprocessing model (Cunningham & Zelazo, 2007) describes evaluation as an iterative process in which stimuli are interpreted and reinterpreted with respect to an increasingly complex set of contextually-relevant representations. At early iterations, evaluations are relatively automatic and particularly rely on stored attitudes, whereas later iterations may be relatively more thoughtful. According to the model, information from sensory cortex about objects in the environment is sent via the thalamus to the amygdala, which is a critical structure for the most rapid and early assessment of hedonic significance. Amygdala activity is sensitive to extremity of valence, even for subliminally presented stimuli (see Cunningham & Zelazo, 2007 for a review). This early evaluation is coarse and simple, but can direct further attention and can initiate preparation for action (via the amygdala's projections to anterior cingulate cortex,

orbitofrontal cortex, the hypothalamus, etc). With increasing iterations, initial evaluations are enriched by information from more sources and feedback, integrating contextual information, one's physiological responses to the stimulus, and so on. Such a process of evaluative integration points to the potential for misattribution at a neurological level. Representations evoked by the US may be among the inputs integrated into an evaluation of the CS. Thus, the iterative reprocessing model, like the literature on feature migration, provides a way of thinking about the mechanism of implicit misattribution that differentiates it from explicit misattribution, that is, the conscious (but mistaken) reasoning about the cause of an event.

3. Implicit misattribution and characteristics of EC

The IMM generates predictions about evaluative conditioning that are sometimes consistent and sometimes inconsistent with empirical research on EC, as is to be expected if prior demonstrations of EC occur through multiple mechanisms including misattribution. Of the previously noted findings that suggest EC has characteristics that differ from PC, many are consistent with the IMM.

First, the IMM is consistent with findings that EC can occur in the absence of contingency awareness. There is simply no reason to assume that noticing and remembering that the CS and US have appeared together would be necessary for the operation of the proposed process. In fact, focusing on the co-occurrence of CS and US could conceivably disrupt implicit misattribution by drawing attention to the actual source of an activated attitude.

Second, whereas the success of PC clearly varies as a function of the consistency of CS-US covariation (signal value), EC can vary as a function of the absolute number of CS-US pairings rather than the regularity with which they co-occur. EC, unlike PC, at least sometimes does not diminish when additional CS-only presentations are added to the CS-US pairings

(Baeyens, Hermans, & Eelen, 1993). Again, this is consistent with the IMM. It is not the CS's signal value in predicting the US that leads it to acquire valence in this case. Instead, it is its co-occurence with the US as a plausible source of an evaluative response that produces EC. Thus, each co-occurrence is an opportunity for valence transfer, and we would expect EC to increase as a function of the number of CS-US pairings (although not, of course, indefinitely).

Third, EC can be resistant to extinction compared to PC. After successful PC, the acquired response to the CS extinguishes if the CS repeatedly appears without the US because such failures of the US to occur diminish the signal value of the CS. Because the basis of the acquired CS attitude according to the IMM is not signal value but instead the very valence previously associated with the US itself, there is no reason to expect that EC should be particularly sensitive to extinction. De Houwer, Thomas, and Baeyens (2001) have reviewed evidence that it is not. Having acquired a particular valence from the US, the CS appearing alone retains that valence and is evaluated differently than before; indeed, the newly acquired attitude is likely to bias subsequent processing of the CS in a manner congruent with the US valence (e.g., Kendrick & Olson, 2009).

IV. Testing the Implicit Misattribution Model

Having noted that the IMM is consistent with many of the properties that have previously distinguished EC from PC, we will now address procedures that can be employed to test it, some unique predictions that it offers, and, hence, more direct evidence for the postulated process. In order to test the IMM, it is ideal to employ EC procedures that could plausibly produce EC through implicit misattribution but are unlikely to do so through other mechanisms. The surveillance procedure would seem to meet this criterion (see also Jones, Fazio, & Olson, 2009). The most pertinent characteristics are the simultaneous presentation of CS and US pairings, the

use of multiple US each paired only once with a CS, and the minimization of contingency and demand awareness. We view simultaneous CS-US presentations as the most crucial methodological key to producing implicit misattribution. The use of multiple US and the minimization of contingency and demand awareness on the other hand function more to inhibit the occurrence of EC through mechanisms other than implicit misattribution, which we will discuss when we describe other mechanisms in greater detail. These characteristics of the surveillance procedure render it suitable for testing hypotheses regarding implicit misattribution.

Though the IMM leads to many predictions about EC, the most distinctive predictions concerning the presence and magnitude of EC effects are united by a concept that we have called "source confusability" (Jones, Fazio, & Olson, 2009). Source confusability refers to any factor of the person, stimuli, or situation that increases the likelihood that the response evoked by the US might be misattributed to the CS. In other words, in some cases it is plausible that a response to the US might be misattributed to the CS, and in others it is not. The simultaneous presentation of CS and US used in the surveillance procedure is conducive to source confusion, for example, because the onset and offset of the CS are identical to those of the US that actually evoked the attitudinal responses. Any factor that increases source confusability should increase the likelihood of EC through misattribution but not necessarily through other mechanisms. With this basic framework, we have conducted multiple tests of the IMM that have provided evidence for its validity.

A. Perceiver characteristics that foster source confusability: Eye gaze shifting

Implicit misattribution is in our view essentially a mistake of automatic perceptual-cognitive processing. The assignment of an internal response to an external cause is made inaccurately. Under what conditions might this be expected to occur? We reasoned that this

should first be particularly likely when the CS and US are processed simultaneously. Specifically, the misattribution of a response to the US to the CS should primarily take place when a mental representation of the CS is active at the same time a response to the US is being experienced. Although we view spatio-temporal contiguity as, by definition, an antecedent of evaluative conditioning, it is implicit misattribution that should be especially contingent upon the degree of overlap between the activation of the US and attention to the CS. Source confusability depends on several variables, and we suspect spatio-temporal contiguity is one of the most important ones. This may be especially true in the laboratory, where the evaluative response to a standard US is not particularly extreme and therefore is likely to be quite transient. It is the necessity that the actual response evoked by a US be experienced during CS processing that leads the misattribution mechanism to be exceptionally dependent on spatio-temporal contiguity.

We reasoned that if EC through misattribution is indeed dependent on this sort of simultaneous processing, then we could measure overt indicators of mental processing in order to predict the extent to which individuals exhibited conditioning. Specifically, we took advantage of the fact that eye gaze is a reasonable outward indicator of internal mental focus (e.g., Kruschke, Kappenman, & Hetrick, 2005). We used an eye tracker, a device that provides a point estimate of visual gaze, to measure patterns of eye gaze while participants were engaged in the surveillance procedure (Jones, Fazio, & Olson, 2009, Study 1). We predicted that the simultaneous processing of CS and US would be associated with one particular pattern of eye gaze. We reasoned that any time that participants' attention shifted between the CS and US an opportunity for misattribution was occurring. A pattern of eye gaze in which a fixation on the US is immediately followed by a fixation on the CS or vice versa (without intervening fixation)

is perhaps the most direct evidence possible that the two stimuli are being mentally processed in close temporal continuity.

The eye gaze data confirmed our expectations. The more frequently a participant spontaneously shifted eye gaze between the CS and US, the more likely that participant was to demonstrate a conditioning effect. The number of separate occasions on which a participant shifted three or more times between the two stimuli during a single 1500 ms trial was especially predictive of conditioning. Such repeated shifting of attention between the two stimuli apparently fostered misattribution. Despite extensive exploratory efforts to identify other correlates of conditioning relating to eye gaze, it was only attentional shifting that predicted the conditioning effect. The eye gaze data were, however, useful for confirming that the participants performed the surveillance task diligently. Participants focused at least once on both the CS and US on the vast majority of trials ($M > 93\%$) and spent, on average, 1350 ms of each 1500 ms trial focused on one of the two objects or in movement. Thus, the CS and US were undeniably processed in close spatio-temporal contiguity, but evaluative conditioning was enhanced among those participants whose patterns of eye-gaze shifting fostered misattribution.

B. **Characteristics of CS-US presentation that foster source confusability**

1. Alternating flashes

Experiments testing the IMM have also involved manipulating factors concerning the presentation of stimuli in the surveillance procedure that could facilitate source confusion. First, we experimentally replicated the correlational finding that attentional shifting fostered EC by causing the CS and US to "flash" back and forth (Jones, Fazio, & Olson, 2009, Study 2). This was achieved by briefly removing each stimulus for an extremely short period of time (25 ms) in alternating fashion such that each stimulus disappeared twice. These brief disappearances

evidently captured the attention of participants. The effect of conditioning was greater in this condition than in a control condition in which CS and US did not flash.

2. Temporal overlap

The IMM also suggests that when misattribution underlies EC, it should be especially sensitive to the degree of spatio-temporal contiguity between CS and US. To compare with PC, consider a hypothetical example in which a neutral tone precedes the occurrence of a negative event, for example the presentation of an unpleasant picture, by several seconds. One might come to dislike that tone as a function of learning that it will lead to a negative event, but one would seem very unlikely to *mistake* that neutral tone as itself causing a negative reaction. Now imagine that the tone and picture occurred simultaneously. It seems much more plausible that misattribution could occur.

Relatively small differences in the temporal elements of stimulus presentation can also prove important in EC. Implicit misattribution is a plausible mechanism underlying EC in one recent experiment that demonstrates this point (Rydell & Jones, 2009). The method bore several similarities to the surveillance procedure in that contingency awareness was avoided, CS appeared with multiple US of a given valence, and CS-US pairings, though not simultaneous, were highly temporally contiguous. The experiment employed a US-competition procedure in which each CS appears with two US per trial. In some conditions, the two US differed in valence: this allows one to examine which US (due to its content or manner of presentation) was more influential in determining the ultimate valence acquired by the CS. In one condition, the onset of the US lasting 2 s preceded the occurrence of the CS, overlapping with it for 250 ms before its offset, at which point the onset of the second 2 s US occurred, overlapping with the CS for 250 ms before the CS offset. In other words, both US partially overlapped with the CS and

did so equivalently. In this condition, evidence for negativity dominance (Rozin & Royzman, 2001) in EC was obtained—negative US were more influential than positive US even though the two sets of US had been equated for extremity of valence. In another condition, one US overlapped with the CS just as described above, but the onset of the second US was delayed such that it was temporally adjacent rather than overlapping with the US. In this condition, the more temporally contiguous US always determined the valence of the CS. This was true regardless of whether a positive or negative US was temporally overlapping—in fact, there was no sign of an effect of US valence here: the effect of temporal overlap completely obscured it, suggesting that even very small differences the temporal characteristics of presentation can be powerful influences on EC.

3. Spatial proximity

A similar argument applies to spatial proximity. Individuals spontaneously localize the spatial source of hedonically relevant events, integrating spatial and affective information (Crawford & Cacioppo, 2002). Therefore spatial proximity is often likely a constraint on implicit misattribution as individuals may have a sense that the object *here* rather than *there* is evoking a given response. Indeed, the visual perception literature has identified proximity as a factor that promotes feature migration (e.g., Cohen & Ivry, 1989; Gallant & Garner, 1988).

Inspired by such reasoning and the findings from the perception literature, we have examined experimentally whether small differences in spatial proximity influence the magnitude of EC in the surveillance procedure (Jones, Fazio, & Olson, 2009, Study 3). We manipulated the spatial proximity of CS and US such that in the "near" condition, the CS and US stimuli (used appeared quite close together. They appeared side-by-side in the middle of the screen, nearly but not quite touching. In the "distant" condition, the CS and US were each again centered along the

vertical dimension but were maximally separated along the horizontal dimension such that each was adjacent to the outer edge of the computer display. As expected, the near condition produced a greater EC effect than did the distant condition. In fact, even though the degree of spatial proximity in the distant condition was still relatively high insofar as both stimuli appeared on the same computer screen separated only by a few inches, this sufficed to prevent a statistically significant conditioning effect from emerging in this condition at all.

4. CS Salience

Though the IMM suggests that CS-US proximity and overlap are particularly important when the mechanism underlying conditioning is misattribution, some degree of CS-US contiguity is a necessary condition for all forms of EC, and CS-US co-occurrence is by definition an antecedent of EC. The IMM is more unique in predicting that the relative salience of the CS and US will serve as a determinant of EC. A basic principle of attribution theory (Heider, 1958), demonstrated in a large number of experiments (e.g., Jones & Nisbett, 1971; Pryor & Kriss, 1977; Storms, 1973; Taylor & Fiske, 1978) is that the more salient or attention-grabbing a potential cause of an event is, the more likely it is this potential cause will be used to explain the event. This leads that to the prediction that when EC occurs through implicit misattribution, if the CS is salient relative to the US, EC should be facilitated, whereas if the US is salient relative to the CS, it should be inhibited. When the CS captures attention, it is likely to be mistakenly identified as the source of an evoked response. When the US captures attention, it is likely to be correctly identified as having evoked that response.

We tested this hypothesis in the surveillance procedure (Jones, Fazio, & Olson, 2009, Study 4) by operationalizing salience as the relative size of the stimulus—larger stimuli being more salient. Typically, in the procedure the CS and US are the same size. In this experiment,

one of the two stimuli was of standard size, but the other was approximately doubled in size, occupying nearly half the computer display. As expected, greater EC was observed when the CS was larger than the US compared to when the opposite relationship was present. No EC effect obtained when the US was larger than the CS. Although this effect makes a good deal of sense from a misattribution perspective, it is also counterintuitive in that it might be assumed that large US would be more rather than less effective in influencing attitudes. We would also note that increasing the relative salience of the CS is probably useful only to a point. Should the CS capture attention too effectively, the US may be neglected and, hence, neither evoke a sufficiently powerful evaluative response to influence the CS nor promote attention shifting between CS and US.

C. Characteristics of CS-US content that foster source confusability

1. Evocativeness

The IMM is also unique in the prediction that the content of the CS and US comprise an important boundary condition of EC due to the potential for content to influence source confusability. Different types of US obviously evoke different types of responses, not all of which could plausibly have originated from the CS. First, we will consider a variable we have termed the "evocativeness" of the US. Evocativeness is a catch-all term for variables like extremity of valence, propensity to produce arousal, and attitude accessibility—variables associated with attitude strength. It is questionable whether powerful US could conceivably evoke responses that are open to misattribution. On the other hand, a more mild and subtle evaluative response to a valenced object could easily be mistaken as having arisen from a complex and potentially meaningful neutral object such as the cartoons, faces, and symbols typically used as CS.

We tested this idea (Jones, Olson, & Fazio, 2009, Study 5) by constructing two sets of US stimuli for use in the surveillance procedure. On the basis of pretesting, we identified one set of US with highly accessible attitudes and another with more moderately accessible attitudes, though in both stimulus sets we selected only US that participants identified by consensus as unambiguously positive or negative. We focused on US attitude accessibility as our operationalization of evocativeness, reasoning that US associated with highly accessible attitudes would evoke responses that are particularly "bound" to their sources by virtue of their strong associative linkage. US associated with less accessible attitudes, on the other hand, would be less evocative and these attitudes might be more susceptible to misattribution. However, we also allowed the two stimuli sets to differ with respect to the natural correlates of attitude accessibility such as attitude extremity because we were interested in testing the general concept of evocativeness as opposed to isolating the contribution of attitude accessibility per se.

The results were interesting. Our hypothesis that mildly evocative US would be more effective at producing EC than strongly evocative US was confirmed, but only after excluding contingency aware participants. An unusually high percentage of participants in the strongly evocative condition (32%) were contingency aware, i.e. could report the valence of US that appeared with one or both CS. This subset of participants also showed quite a strong EC effect; but no EC effect obtained for the contingency unaware participants in the strongly evocative condition, despite this statistical test having much greater power due to this subset of participants outnumbering the contingency aware participants 2:1. Meanwhile, only three participants (5%) exhibited contingency awareness in the mildly evocative condition, and these participants were excluded from further analyses. The remaining contingency unaware participants in this case, however, did in fact exhibit a conditioning effect. Thus, in addition to confirming that mildly

evocative US can be more effective in producing EC than strongly evocative US (another clearly counterintuitive finding that is generated by the IMM) at least in the absence of contingency awareness, this experiment suggested that different mechanisms of EC may operate as a function of the specific CS-US content. Mildly evocative US were conducive to producing the misattribution of a response to the US to a CS, but strongly evocative US operated through a different mechanism that would appear to be mediated by contingency awareness.

The finding regarding attitude accessibility is also consistent with the phenomenon of US pre-exposure, in which presentations of the US alone before the conditioning phase reduce subsequent conditioning. US pre-exposure, especially if participants repeatedly evaluated the US, would have the effect of increasing US attitude accessibility (Powell & Fazio, 1984) thus potentially reducing implicit misattribution by increasing evocativeness[3].

2. Feature matching

Another characteristic of CS-US content related to source confusability is feature matching or belongingness. The idea here is that in order for misattribution to be feasible, some minimal degree of fit or appropriateness between the response evoked by the US and the CS serves as another boundary condition. In fact, evocativeness can be considered a particular case of feature matching. Let us return to Zanna et al.'s (1970) experiment as an illustration of feature matching. In that study words (e.g. "light" and "dark") served as CS and a painful electric shocks served as US. Could misattribution explain EC here? It seems doubtful because one could not possibly mistake, implicitly or otherwise, the response to the painful shocks as having come from the words. Such a US seems both too extreme and too incompatible in its

[3] Hammerl, Bloch, and Silverthorne (1997) demonstrated a US pre-exposure effect and explained it in terms of habituation to the US. That is, EC was reduced because the now blasé US no longer evoked a sufficiently positive or negative response to influence the CS. This explanation is also plausible and indeed, because the procedure used by Hammerl and colleagues was not in our view likely to have operated through misattribution, it is probably a better explanation for that experiment.

specific content for this to occur. Those words, on the other hand, can easily function as a signal that the US will occur. No particular fit or appropriateness between the signal and the valenced event are necessary for the presence of signal value. Though feature matching might conceivably ease signal learning, it would not appear to be a boundary condition in the same way that it is for misattribution. Though feature matching has received fairly little attention in the EC literature, some studies are consistent with it. For example, one experiment demonstrating cross-modal conditioning between CS faces and US smells found transfer of valence between US and CS only when the US smells were "plausibly human" (Todrank, Byrnes, Wrzesniewski, & Rozin, 1995). Consumer psychological research on EC similarly has demonstrated an effect of feature matching such that celebrity endorsers are more effective when fit between the celebrity and product is high (Till, Stanley, & Priluck, 2008). The role of feature matching and its relation to different mechanisms of conditioning is a worthy topic for future research.

V. Mechanisms of Evaluative Conditioning

The previous section reviewed research suggesting that implicit misattribution can produce EC. We do not, however, argue that implicit misattribution is synonymous with EC. Rather, it is one of multiple mechanisms that can intervene between the antecedent of spatio-temporal co-occurrence between CS and US and the consequence of attitude change toward the CS.

Thus, we agree with De Houwer (2007), who has provided an elaborate justification for this position, that EC is best conceptualized as an effect. The primary alternative is conceptualizing EC as a particular process. De Houwer (2007) argues that doing so is problematic for multiple reasons. First, it renders it difficult to determine whether EC has occurred or not because EC as a theoretical process would not be directly observable. It further

leads to contention regarding whether a successful conditioning procedure is producing "true" EC. We would also note that, whether by intention or not, the term has de facto been used historically to refer to a variety of processes that resemble one another due only to their sharing the antecedent of spatio-temporal co-occurrence and the consequence of attitude formation or change. Conceptualizing EC as a particular process, thus rendering many prior uses of the term improper, would be quite confusing for readers of this literature and thus should require quite compelling justification. De Houwer (2007) also argues the view of EC as particular theoretical process hinders theoretical development by stifling new theorizing about process and by treating the "how" question as closed at an inappropriately early juncture. The primary criticism of conceptualizing EC as an effect, so far as we can tell, is that its inclusiveness might be viewed as detrimental. It leads various processes with quite different characteristics to be grouped together. This is a reasonable concern. Accepting that multiple mechanisms can underlie EC does, for example, often make it difficult to make simple assertions about EC without qualification. On the other hand, the current status quo in which it is common to pronounce definitively that EC has one characteristic or another, as though it is invariant and consensually understood, is also problematic. It interferes with effective communication and is nearly always subject to empirical counterexample.

Ultimately, this question of conceptualization is largely semantic, and consensus regarding it, although preferable, need not be the ultimate priority. Whether a researcher chooses to call one or many processes "evaluative conditioning" is less important than whether that researcher clearly specifies the proposed process that underlies hypothesizing and theorizing and employs methods that are suitable to generating that process. It is quite problematic, on the other

hand, to employ a conditioning procedure while remaining agnostic to the underlying process being studied and yet to then generalize broadly about the results' implications for EC.

A recent meta-analysis of evaluative conditioning (Hofmann, De Houwer, Perugini, Baeyens, & Crombez, in press) illustrates our very reservations about broad generalizations regarding characteristics of EC. In our view, this review paper fails to take seriously the problem of multiple mechanisms. Though the authors do not claim a single mechanism operates in all EC studies, the literature review is structured such that multiple models of conditioning are presented as competing to explain the entirety of the body of EC literature. Unsurprisingly, all of the reviewed models of EC are found to be incompatible with at least one major finding of the meta-analysis, though the propositional account (see below for an explanation) fares relatively well. Of much relevance here, the IMM appears incompatible with some of the meta-analytic findings. We would point out, however, that the vast majority of experiments in the EC literature are unlikely to have operated through implicit misattribution. The most frequently used procedure is some variant of the picture-picture procedure (Levey & Martin, 1975), which is different than the surveillance procedure in a number of ways, including that CS-US presentations are not simultaneous, that a given CS is paired repeatedly with a single US rather than multiple US, and that a cover story and task are not employed to minimize contingency awareness. Had the surveillance procedure first been published and popularized in 1975 and the picture-picture procedure in 2001, like the surveillance procedure, a meta-analysis of EC would quite likely look radically different. The characteristics of EC are highly dependent on the method employed, yet researchers are only beginning to understand this relationship. It is problematic to lump all findings regardless of procedural variations, and then to strongly interpret the modal finding as a universal. The dominance of particular procedures is largely a function of historical precedent.

Accounts of the characteristics of EC must be regarded as conditional upon the mechanism that any given procedural variation seeks to promote.

Keeping in mind the strong likelihood that previous demonstrations of EC have operated through a variety of cognitive mechanisms, we now turn our attention to a further comparison of various theoretical mechanisms that have been proposed to underlie EC.

A. Implicit misattribution

We have already discussed implicit misattribution as a mechanism in some detail. As a reminder of some of its most important characteristics, implicit misattribution is acquired as a function of the absolute number of CS-US pairings—each pairing acts as an opportunity for misattribution, which likely act cumulatively. It does not require contingency awareness, which could even interfere with the process, especially if noting the co-occurrence leads one to correctly identify the source of the response to the US. Our findings in the standard surveillance procedure are typically that contingency awareness, to the extent that it occurs, and EC are uncorrelated. The consequence of EC through implicit misattribution is that the conditioned attitude is based on the transfer of affect from the US to the CS. Thus, there is no reason that presentations of the CS without the US would produce extinction.

B. Pavlovian Classical Conditioning

Staats and Staats' (1958) original demonstration of EC was titled "Attitudes established through classical conditioning." They assumed, apparently, that the process underlying the effect they demonstrated was an attitudinal version of Pavlovian classical conditioning (PC). Thus, the terms "conditioned stimulus" and "unconditioned stimulus" were borrowed from that literature. As we explained at the beginning of this paper, we think it's unlikely that PC operated in the Staats and Staats procedure. Nevertheless, we recognize that the signal learning at the

heart of PC is also a valid mechanism of EC, and we argued earlier that it is likely to have occurred in the experiment by Zanna and colleagues. The proposition that EC is PC has been strongly advocated by some; see Field (2000a, 2000b) as excellent examples of the position.

A number of studies have demonstrated that signal learning (PC) can be accompanied by changes in evaluation of the signal (EC). For example, Hermans, Vansteenwegen, Crombez, Baeyens, and Eelen (2002) used a neutral human face as an indicator that a painful electric shock would follow. Participants learned this contingency very well and reported a high degree of certainty about it. Furthermore, when attitudes towards that face were measured with an evaluative priming task, they were evaluated much more negatively than a control face that had been judged as equivalent before conditioning. Purkis and Lipp (2001) also found that pictures followed by shocks were subject to both signal learning (as measured by skin conductance response) and evaluative learning, though both effects were limited to those who could report the CS-US contingency.

When changes in liking for the CS are a byproduct of PC, the basis of the formed or changed CS attitude is its signal value in predicting the occurrence of the US. In this case, attitude acquisition is a function of signal learning, which is highly dependent if not completely contingent on awareness of the relation between the two objects (see Lovibond & Shanks, 2002). It will occur most readily when the CS invariably and immediately signals occurrence of the US, i.e., under conditions of statistical regularity. According to the PC perspective, if the CS ceases to signal the US, it will lose its signal value and the conditioned attitude will extinguish. However, the existing data are ambiguous on this point. Lipp, Oughton, and LeLievre (2003) found extinction of conditioned attitudes at the end of the conditioning phase, but also found that a difference in evaluation of the CS+ and CS- re-emerged in a post-test. Hermans, Crombez,

Vansteenwegen, and Eelen (2002) and Blechert, Michael, Williams, Purkis, and Wilhelm (2008) found extinction on at least some measures of expectancy regarding the occurrence of the US but resistance to extinction on measures of evaluation of the CS. All of these studies involved traditional PC methodologies. It may be the case that the attitudinal consequences of signal learning are more robust to extinction than its physiological consequences (e.g., skin conductance) or effects on explicitly-measured expectancies. Perhaps merely remembering that an object did reliably signal a hedonic event at an earlier point in time is sufficient reason to continue to judge it as good or bad.

For an excellent demonstration of the functioning of EC occurring through signal learning, see Dawson, Rissling, Schell, and Wilcox (2007). In this study, the CS (initially neutral faces) always preceded specific and repeating positive (a cute baby) or negative (an injured woman) US. Skin conductance responses, startle eyeblink responses, unpleasant-pleasant scale ratings, and an evaluative priming measure indicated affective and physiological consequences of those pairings, but each of these measures only indicated effects of the pairings in those who could explicitly report contingency awareness. Further, extinction CS-only trials quickly eliminated effects on the physiological measures administered during this phase, the eyeblink and skin conductance responses, but self-report and priming measures of attitude were not reassessed. On the basis of the findings noted earlier, we would speculate that such evaluative measures may have shown more persistence.

C. Belief formation (stimulus-attribute association)

Another means through which pairing two objects might facilitate EC is if that pairings leads one to associate a specific attribute of the US with the CS. Fishbein and Middlestadt (1995) have gone so far as to claim that all attitude change, including evaluative conditioning, is

mediated by attribute-based cognitions about the attitude object as detailed in the expectancy-value model (Ajzen & Fishbein, 1980). Though we obviously disagree that this single mechanism adequately describes all attitude change, or even all evaluative conditioning, we certainly concur that it can be considered *a* mechanism.

For example, Kim, Allen, and Kardes (1996) paired the CS "Brand L Pizza" with US such as runners and a racecar. Consequently, participants both came to like the brand more and came to believe it was more likely that Brand L Pizza provided rapid delivery. Notably, this was particularly true for contingency aware participants. Further, the conditioning procedure's effect on attitudes was mediated by the beliefs formed about speed of service. A second experiment, however, employed various US and found evidence for dual mediation of such EC both by belief formation and by a non-cognitive mechanism independent of the attribute paired with the CS. Thus, the authors argued against the position that EC only occurs through belief formation.

These studies are reminiscent of the phenomenon of spontaneous trait transference, wherein a communicator describes a person as possessing a particular attribute and the communicator comes to be seen as possessing that attribute (Skowronski, Carlston, & Mae, 1998). This effect is based on simple object-attribute associations rather than being attributional (Carlston & Skowronski, 2005) and would appear to be a very interesting example of EC.

When EC occurs through object-attribute association, the conditioned attitude toward the CS is based on an association with a valenced attribute possessed by the US. This process would seem to be facilitated by feature matching or fit between the CS and US, though perhaps not so sensitive to it as misattribution. Belief formation appears to be quite dependent on contingency awareness. These are tentative suggestions; however, the functioning of this mechanism should be further studied in EC procedures. Research exploring this potential mechanism is much more

limited than, for example, that of Pavlovian conditioning. Much as Pavlovian conditioning theory provided a theoretical framework for those studies, expectancy-value frameworks of attitudes can guide thinking about EC through belief formation.

The earlier discussions of feature migration and implicit misattribution might make one wonder if implicit misattribution of particular attributes rather than a more generalized evaluative response could constitute another form of EC. The available evidence regarding EC through object-attribute association, as noted, displays characteristics more consistent with a more thoughtful, cognitive mechanism. We are not aware of any particular evidence for attribute misattribution influencing evaluation, but it would appear to be a possibility.

Both the belief formation and Pavlovian conditioning mechanisms qualify as models of propositional learning (see De Houwer, 2009). The first critical feature of propositional learning is that the linkages between mental representations go beyond mere association in memory capable of producing activation or inhibition. Instead, these linkages possess specific content, propositions that differ from simple associations in that they suggest a truth value (Strack & Deutsch, 2004). Pavlovian conditioning thus involves learning the proposition that the CS *signals* or potentially that it *causes* the US. Such propositions can clearly be true or false in a way that mere association cannot. Belief formation involves the proposition that the CS *has the characteristic of* the US or some element thereof. The second critical feature of propositional learning is that it is non-automatic. “Proponents of propositional models describe these processes as a form of conscious, effortful, time consuming, and controlled reasoning” (De Houwer, 2009, p. 27). This non-automaticity is consistent, for example, with the observation that EC can depend on contingency awareness, or that it can consume attentional resources (Pleyers, Corneille, Yzerbyt, & Luminet, 2009). Though mechanisms involving propositional learning

may differ as a function of the type of proposition involved (e.g., forming and testing the proposition that the CS *predicts* the US differs from proposing that the CS *has a characteristic* related to the US), because they are non-automatic they will also share many characteristics and might be considered a class of mechanisms that have more in common with one another than with implicit misattribution or the non-propositional mechanisms summarized below.

D. Stimulus-stimulus association

As opposed to the propositional learning mechanisms just reviewed, the stimulus-stimulus model is the classic example of simple, bottom-up, automatic associative learning. EC may occur through the development of a specific stimulus-stimulus association, that is, an association in memory between a CS and a representation of the US as a whole as opposed to some attribute thereof. In this model, basic, low-level associations form between the CS and US. When the CS is later encountered, its evaluation is influenced by the specific US that it brings to mind. A CS that reminds one of a positively-valued US will be viewed more favorably than one that calls to mind a negatively-valued US. This model has been discussed most extensively by Gawronski and Bodenhausen (2006).

The hallmark of a procedure that could operate through this mechanism is the repeated pairing of a CS with a particular US, as opposed to several non-repeating US. This will create CS-US associations, which would not necessarily *require* contingency awareness but would certainly be enhanced by it. In a later section on contingency awareness, in fact, we will describe a study suggesting that in the case of CS-US association mediated EC, the truth is indeed more nuanced than *does* or *does not* require contingency awareness. That is, the role of contingency awareness may depend on one's familiarity with the CS (Ruys & Stapel, 2009). Similar arguments might be made for the relevance of feature matching to this mechanism.

Particularly strong evidence for the viability of this account comes from the phenomenon of US-revaluation. According to this account, the basis of the conditioned attitude is an indirect association with the US attitude. Therefore, if that latter attitude changes (i.e., US-revaluation), then the post-conditioning CS attitude should also change. Indeed, such an effect has been carefully demonstrated by Walther, Gawronski, Blank, and Langer (2009). US-revaluation can have practical consequences. For example, imagine that EC is achieved by associating a CS brand with a well-liked US celebrity. Should that celebrity become involved in a scandal and fall out of favor, the CS would become less liked as well.

It appears that CS-US associations can influence evaluations of not just the CS but also objects that are (pre)associated with the CS—the spreading attitude effect (Walther, 2002). For example, imagine that an individual strongly associates Bob and Karen, who are often together. If that individual's attitude toward Karen changes, the individual's attitude toward Bob would change similarly as well. The extent of such spreading of activation likely depends on the strength of association and number of links over which such activation is spread (i.e., the "fan" effect). Walther (2002) also provided evidence that the spreading attitude effect is resistant to extinction. This likely holds for EC that is produced by stimulus-stimulus association generally—there is little reason to expect that CS-only presentations would strongly undermine the basis of the acquired attitude.

E. Conceptual recategorization

Finally, Davey (1994) has argued that EC may arise from *conceptual recategorizaton*, wherein the presence of the US highlights conceptually congruent features of the CS, leading one to categorize the CS in a more positive or negative way than one otherwise would have. This is essentially a priming mechanism wherein the US serves as a prime directing attention to

congruent elements of the CS. Under what conditions this sort of priming will lastingly influence CS attitudes is not clear. This mechanism has received little theoretical attention and even less empirical scrutiny. It is interesting in that it posits a basis for the conditioned attitude that is entirely different from those mechanisms previously discussed—the increased salience of particular features of the CS itself. A similar but conceptually distinct mechanism that would produce this same result would involve the US valence causing biased processing of the CS features.

VI. Implications for Evaluative Conditioning

Having described the IMM in its context as one of multiple mechanisms that can intervene between CS-US co-occurrence and CS attitude change, we will now address its implications for EC generally, as well as remark at times upon how other mechanisms relate to these same issues.

A. Elusiveness of evaluative conditioning

We earlier remarked on the apparent consensus in the world of EC research that the phenomenon is relatively elusive. It is common for the phenomenon to fail to obtain for seemingly inexplicable reasons, indicating that its boundary conditions are poorly understood. In the case of implicit misattribution, one boundary condition we would like to highlight is that spatio-temporal contiguity is a necessary but *not* a sufficient condition to produce EC. In the experiments reported by Jones, Fazio, and Olson (2009) each of the control conditions not fostering source confusion certainly still entailed spatio-temporal contiguity. However, those conditions (e.g., CS and US positioned distantly on the computer screen, or images that involved a US of relatively greater salience than the CS) tended not to produce EC. It appears, then, that source confusability is a very pertinent boundary condition for implicit misattribution.

Whether or not one accepts that EC is best considered an effect that manifests through multiple mechanisms, it now seems beyond question that the various procedures that have been described as EC procedures do not share a common mechanism. We suggest that a major source of conditioning failures may be the combination of elements of multiple procedures that appear similar at a surface level but function differently. For example, sequential, non-overlapping CS-US pairings can produce EC, as is typical in the commonly used picture-picture paradigm (Levey & Martin, 1975). Both PC and stimulus-stimulus association are likely mechanisms here; implicit misattribution, though not theoretically impossible, is at the outset rendered unlikely because of the lack of CS-US temporal overlap. It is also possible to produce EC by pairing a single CS with multiple US (e.g., Olson & Fazio, 2001). However, combining a CS *sequentially* with *multiple US* is non-optimal with respect to the major mechanisms of conditioning. The misattribution mechanism favors the simultaneous rather than the sequential pairing of CS and US. PC and stimulus-stimulus association are also relatively unlikely when a CS is paired sequentially with multiple US. With respect to PC, learning that the CS signals an abstract type of event (good or bad) is harder than learning it signals a specific US. With respect to stimulus-stimulus association, various CS-US associations produced by one-trial learning would have to underlie EC rather than a single CS-US association strengthened by repeated pairing.

These outcomes are not impossible; they may be rather more feasible, for example, if very strong US are employed, but the point is that procedures are not being optimized to operate through a particular mechanism. Sweldens (2009) recently addressed this issue of procedural and mechanistic compatibility. He compared sequential pairings of a CS and multiple US to sequential pairings of a CS with a single US, as well as to both sorts of simultaneous pairings. Only the sequential pairings with multiple US failed to produce a statistically significant EC

effect. The essential point here is that producing a robust EC effect requires optimization of the particular parameters that are critical to the specific mechanism one is attempting to promote. Is it signal-learning, belief formation, stimulus-stimulus association, implicit misattribution, or whatever? Each process necessitates attention to different concerns.

B. Measurement of EC with explicit and implicit measures of attitude

Explicit measures of attitude are those standard means of assessing attitude involving introspective self-report. Implicit measures attempt to assess attitudes indirectly through various methods, often relying on response time measurements. It is not uncommon for explicit and implicit measures of attitude to diverge for multiple reasons (see, e.g., Fazio & Olson, 2003; Gawronski & Bodenhausen, 2006; Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2005; Payne, Burkley, & Stokes, 2008). Consideration of the mechanism underlying EC provides some insight into which are most likely to manifest attitude change and why. For a detailed discussion of this issue and EC from the perspective of a stimulus-stimulus model, see Gawronski and Bodenhausen (2006).

With respect to implicit misattribution, the basis of the conditioned attitude is the CS's association with an affect transferred from the US. From the perspective of the MODE model (Olson & Fazio, 2009), implicit measures primarily reflect one's automatically activated attitude. This is because by virtue of their indirect nature, they are not subject to reflective evaluation. The MODE model posits that judgments and behaviors are by default spontaneously guided by automatically activated attitudes. However, if one has both the motivation and opportunity to deliberate further about the judgment or behavior, such deliberations may produce a different outcome than would follow from the automatically activated attitude. Therefore, EC through implicit misattribution should be expected, if successful, to always manifest on implicit

measures. They will also manifest on explicit measures in some circumstances, but if one has the time and inclination to further evaluate the CS, they may not. When thoughtfully considering the CS, one may not rely on the activated attitude for various reasons, including that it may not seem justified. Very simple CS lacking meaningful features, in particular, may be reported as neutral despite evoking some evaluative response. This sort of concern motivated the use of the forced-choice dependent variable in Jones, Fazio, and Olson (2009), which circumvents the problem by requiring participants only to find some reason to (not) prefer the CS to another object.

We recently tested the notion that attitudes formed implicitly through EC appear primarily as "gut-level," affective responses (Kendrick & Olson, 2007). In one such study, participants underwent the surveillance procedure prior to completing some personality measures about which they were given false feedback. Some participants were told the personality tests indicated that they had excellent intuition and should "trust [their] gut" when making decisions, while others were told that they had superior rational skills and should "trust [their] reasoning." When all participants later evaluated the CS and filler items, only those told to "trust their gut" evidenced an EC effect. In a second study, participants evaluated several cartoon characters of the same genre as the CS after undergoing conditioning and prior to reporting their evaluations of the actual CS. Again, they were given false feedback; some were told they were quite skilled at judging the characters whereas others were told they were not. Only the former participants showed an EC effect. Thus, it appears that believing one is an expert, and, hence, entitled to judge, increases reliance on intuitive processes when expressing judgments (see also Croizet & Fiske, 2000).

In related work, Gawronski and LeBel (2008) reported an experiment in which EC was evident on an implicit but not an explicit attitude measure. In this case, participants were instructed to focus on the reasons for their CS evaluations, increasing the motivation to deliberate about them. However, in another condition, participants were instructed to focus on their feelings while evaluating the CS. This gave participants license to rely on their automatically activated attitudes, and EC was again evident on an implicit attitude measure but in this case also on an explicit attitude measure. Similar findings have recently been reported by Grumm, Nestler, and von Collani (2009) in EC experiments in which one's self served as the CS and implicit and explicit measures of self-esteem were employed.

Whitfield and Jordan (2009) have examined the interplay of associative and propositional attitudinal processes. They gave participants valenced information about novel individuals, which was intended to produce propositional learning. Attitude change was evident on both implicit and explicit attitude measures, but mediational analyses indicated that the propositional learning directly influenced only explicitly reported attitudes, and scores on the explicit measure mediated change on the implicit measure. On the other hand, other experiments involved EC in a variant of our standard surveillance procedure. In this case, attitude change was also evident on both measures. This time, however, mediational analyses indicated that EC (likely through implicit misattribution) directly affected change on the implicit measure, and scores on the implicit measure mediated change on the explicit measure. These findings further highlight the differences previously discussed between mechanisms of EC.

C. Contingency awareness

Debate over the role of contingency awareness in EC is longstanding (see Field, 2000a; De Houwer, Thomas, & Baeyens, 2001 for reviews). For the most part, we think, the issue can

be considered resolved. One can argue that EC occurs exclusively when one is contingency aware (e.g., Field, 2000a; Lovibond & Shanks, 2002; Pleyers, Corneille, Luminet, & Yzerbyt, 2007) and be correct insofar as "EC" signifies a particular process of propositional learning such as PC or belief formation. Many, many studies have found just this. When EC is occurring through propositional learning, contingency awareness is vital because it is the CS-US co-occurrence itself that provides the evidentiary basis for the propositional reasoning that produces EC. At the very least, some transient sort of contingency awareness seems necessary. More generally, increased contingency awareness will increase EC operating through these mechanisms.

At this point, however, it cannot reasonably be denied that one or more processes can also intervene between spatio-temporal co-occurrence and attitude change that are not dependent upon contingency awareness. Though sundry methodological criticisms have been raised to cast doubt on findings of EC obtained in the absence of awareness, the finding continues to emerge in studies that take efforts to rule out each concern as it is raised (for examples, see Olson & Fazio, 2001, 2002; Sweldens, 2009; Walther & Nagengast, 2006). Many, many studies have also found that EC can occur in the absence of contingency awareness, and even that contingency awareness can disrupt EC (e.g. Fulcher & Hammerl, 2001). One can also argue that EC is not dependent upon contingency awareness and be correct insofar as "EC" means a non-propositional mechanism such as implicit misattribution. Further complicating matters, there is no reason to assume that awareness-dependent and awareness-independent mechanisms cannot operate simultaneously to produce EC within a single procedure. Indeed, Kim, Allen, and Kardes (1996) have provided evidence of exactly such dual mediation.

Particularly compelling evidence for EC in the absence of contingency awareness comes from demonstrations of EC when the CS, US, or both are presented subliminally. Such experiments (e.g. De Houwer, Baeyens, & Eelen, 1994; De Houwer, Hendrickx, & Baeyens, 1997; Dijksterhuis, 2006; Gawronski & LeBel, 2008; Krosnick, Betz, Jussim, & Lynn, 1992; Rydell, McConnell, Mackie, & Strain, 2006) produce EC that cannot be dependent on contingency awareness because participants are not aware that pairings are occurring at all. Some, however, have viewed these findings with skepticism (see Pleyers, et al, 2007; Stahl, Unkelbach, & Corneille, 2009) on various methodological grounds. For example, some studies lacked a test of whether the "subliminal" stimuli actually went undetected. In others, US valence was manipulated between-subjects rather than within-subjects, raising concerns about potential mood artifacts. In others, the CS was a familiar and multifaceted object (such as the self), and EC effects may have reflected temporary activation of particular aspects of the object rather than attitude *change* per se. While some of these criticisms suffice to raise doubt about particular studies, they do not in our view persuasively justify the dismissal of the whole literature on subliminal EC. For example, a set of five studies (from De Houwer, Baeyens, & Eelen, 1994; and De Houwer, Hendrickx, & Baeyens, 1997), though apparently acceptable methodologically, has been dismissed by critics for failing to consistently obtain EC. However, such an effect was obtained in three of five reported experiments, and a meta-analysis of these studies (De Houwer, Hendrickx, & Baeyens, 1997) indicated a small but significant conditioning effect. Given the well-known fragility of EC generally, a couple of individual failures to obtain the effect does not seem to be a compelling reason to dismiss the overall conclusion that EC can occur when both CS and US are not presented supraliminally.

We would also like to emphasize some of the evidence for subliminal EC that appears to have been overlooked in prior discussions. Kawakami, Phills, Steele, and Dovidio (2007, Study 2) employed "push/pull" conditioning, in which the positive US and negative US are approach and avoidance motions enacted with a joystick. Each movement was preceded by a subliminal prime of a White or Black face. Even though all participants included in analyses were unaware of any faces appearing in the task, those for whom Black was paired with approach and White with avoid showed attenuated racial bias on an implicit association test relative to controls. Similar push/pull effects with subliminal CS have been obtained by Jefferis and Fazio (2008) and Jefferis, Loersch, and Fazio (2007). For example, Jefferis et al. (2007) paired subliminally-presented non-word CS ("nimono" or "dupate") with approach and avoidance joystick movements and observed EC. That novel non-words were employed, CS were counterbalanced, tests of EC were conducted within-participants, and that the CS were presented subliminally (for only a single 13 ms screen refreshment cycle followed by a mask) are all important given the methodological criticisms of other experiments that have been raised. In sum, these push/pull studies appear to provide strong evidence for subliminal EC.

Thus, we consider the question of whether EC can occur in the absence of awareness to be resolved. It can. However, this is not to say that researchers should not continue to consider questions about contingency awareness. Nor is it to say that recent research on contingency awareness has not been valuable. Recent developments in the measurement and conceptualization of contingency awareness, for example, are important. First, Pleyers and colleagues (2007) have correctly pointed out that a typical way of measuring contingency awareness is flawed. Sometimes, researchers will conduct participant-based analyses of contingency awareness leading to the categorization of a *participant* as aware or unaware as a

function of the response to a majority of the CS. Such analyses can falsely suggest EC in the absence of awareness when in fact EC is only occurring for that subset of CS for which contingency awareness is present. Thus, item-based analyses examining contingency awareness per CS are preferable. However, Sweldens (2009) has conducted item-based analyses and found EC in the absence of contingency awareness under certain conditions (specifically, when CS-US pairings were simultaneous and each CS appeared with various US, as in the surveillance procedure). Sweldens's findings provide further corroboration of the inferences that we have drawn from our own use of a recognition measure assessing awareness of the CS-US pairings. As noted earlier, multiple experiments have failed to reveal better than chance recognition accuracy.

An additional recent conceptual development regarding contingency awareness also deserves mention. Stahl, Unkelbach, and Corneille (2009) distinguished between *identity awareness* and *valence awareness*. The former refers to the ability to identify the specific content of the US that appeared with a CS, whereas valence awareness refers to the ability to identify generally which valence of US appeared with a CS. They argued that whereas identity awareness is typically assessed, valence awareness may be more critical. They found that EC was only evident in valence aware participants and that identity awareness did not tend to contribute to EC beyond valence awareness. Therefore, some findings that EC occurred in the absence of awareness may be mistaken because identity rather than valence awareness was assessed. Although this may be the case, it is at odds with our research using the surveillance procedure. Our funneled debriefing measures are valence awareness measures. A small minority of participants in each study report valence awareness, despite our efforts to prevent it, and are excluded from analyses. Moreover, the recognition test that we commonly employ

involves both identity and valence awareness, because it compares confidence that actual CS-US pairings were presented to confidence that the CS appeared with US of the opposite valence from the actual pairings.

The discrepancy between our results and those of Stahl, Unkelbach, and Corneille may arise because implicit misattribution could not operate in their procedures. Although the four experiments they reported shared several procedural similarities with the surveillance procedure, it is not clear whether the pronounceable non-words or consumer items used as CS and the picture US (their content was not reported) were conducive to source confusion. Also, when each CS was paired with multiple non-repeating US as in the surveillance procedure (Experiments 1 & 2), the number of distinct US (5) paired with each CS was much lower than in the surveillance procedure (20). Thus, there simply may not have been sufficient trials to produce reliable EC in the absence of awareness. Also, little effort was made to prevent contingency awareness, resulting in much awareness and reducing the power of analyses on the contingency unaware subset of the data. Regardless, the distinction between identity and valence awareness is a valuable one and should be the topic of further research.

It will also be useful to continue to refine the scientific understanding of exactly how contingency awareness relates to each mechanism of EC. We mentioned earlier that recent research has suggested that the relation between contingency awareness and the stimulus-stimulus mechanism is complex. Ruys and Stapel (2009) argued that when EC occurs through CS-US associations, the role of contingency awareness should depend on whether the CS is novel or familiar. Novel objects have few associations in memory with other objects. Thus, after it has earlier been paired with a US, when the CS is encountered the US is relatively certain to be activated. Further, little is activated other than the US and its associated attitude; thus they

are relatively likely to influence judgments of the CS. On the other hand, familiar objects have many associations. Following conditioning, the CS-US association is one of many linkages that must compete for spreading activation when the CS is activated. If many associations are activated, the likelihood that the US linkage will in particular guide evaluation is smaller. Therefore, they argued, EC with familiar CS is unlikely unless the CS-US association is especially strong. One way this link might be particularly strong is if one is contingency aware. Consistent with this hypothesis, Ruys and Stapel (2009) observed EC in the absence of contingency awareness for a novel CS but only with contingency awareness for a familiar CS.

Although contingency awareness appeared to enhance EC in Ruys and Stapel's work, as well as in the Jones et al. (2009) experiment involving strongly evocative US (and in many others as well), we have also reviewed evidence that contingency awareness can deter EC (e.g., Fulcher & Hammerl, 2001; Walther & Nagengast, 2006). Theories of social judgment have long extolled the view that social perceivers often attempt to correct for the presence of perceived bias (e.g., Wegener & Petty, 1995; Wilson & Brekke, 1994). In the context of EC, a perceiver may take note of the potential biasing impact of a US on their judgments of a CS, and wish to correct for that influence Such a corrective process may have been operating in these experiments. Hence, contingency awareness need not facilitate EC, and future work should investigate when contingency awareness encourages versus discourages EC. Which it does may depend on whether individuals view the contingencies as an unwelcome source of mental contamination that they are then motivated to correct, or as a valid source of information that they can willingly accept as a basis for their evaluative judgments.

D. Can conditioning procedures influence dimensions other than valence?

We have argued that awareness of contingencies is unnecessary for (and, indeed, may create a barrier to) EC produced via the misattribution mechanism. However, from some of our own and others' recent work, it appears that some categorization along the to-be-learned dimension is necessary for misattribution to occur. Years of research indicates that accessible attitudes are activated automatically upon exposure to a stimulus (Fazio, Sanbonmatsu, Powell, & Kardes, 1986). Assessing the valence of stimuli in one's environment is critical to the pursuit of pleasure and the avoidance of pain, and few other dimensions of categorization can be argued to have similar significance. It is this primacy of the evaluative dimension that we argue promotes activation of the attitude associated with the US, source confusion and hence, implicit EC.

Recent research suggests that implicit covariation learning is less likely among dimensions other than valence. For instance, Meersman, De Houwer, Baeyens, Randell, and Eelen (2005) paired images of gender-ambiguous infants (or Japanese names) with clearly gendered infants (or English names), and saw no evidence of transfer of gender among participants who were unaware of the contingencies. Similarly, using our own paradigm we saw no evidence of covariation learning when one of our usual CS was consistently paired with large/fast stimuli and the other with small/slow stimuli (Experiment 1a/b; Olson, Kendrick, & Fazio, 2009). That gender-ambiguous infants, Japanese names, or cartoon characters were not seen as gendered, small or large, or slow or fast among contingency-unaware participants after being repeatedly paired with such stimuli is consistent with our reasoning that such dimensions are less likely to be utilized for categorization, and hence, do not lend themselves readily to implicit learning via source confusion.

We provided more direct evidence that categorization along the to-be-learned dimension plays a critical role in implicit learning in a second experiment (Olson et al., 2009). Focusing on the size dimension, some participants in this study were exposed to several subliminal presentations of size-related words (e.g., *huge, teeny*) prior to undergoing the conditioning procedure. For these participants, the size dimension had been made more accessible, hence increasing the likelihood that it would attract attention and that stimuli would be categorized according to it (see Smith, Fazio, & Cejka, 1996). Such US as an ocean liner and a hippo would be spontaneously categorized as large, whereas US such as an ant and a button would be categorized as small. As expected, participants in this condition evidenced implicit transfer of the size dimension to the CS; in a "size estimation task" that followed the conditioning procedure, the CS paired with small items was estimated by participants to be smaller than the CS paired with large items, despite the participants' lack of awareness of the pairings. Non-primed participants evidenced no conditioning. Hence, attention to the to-be-learned dimension appears necessary for implicit misattribution to occur. Attitudes toward US are automatically activated and subject to potential source confusion; in contrast, attributes such as size need not necessarily be noted, but can be encouraged by relevant mindsets.

E. EC as Intervention

EC experiments are nearly always conducted with neutral CS. In fact, it is sometimes defined as involving pairing neutral objects per se with valenced objects. However, it is an open question to what extent EC procedures can be effective when positive or negative CS are paired with US of opposite valence. Regarding the implicit misattribution mechanism, it might seem that it would be particularly difficult to achieve such an effect because the opposing valences of the CS and US reduce source confusability. However, we would like to point out that many

positive and negative attitudes involve some degree of ambivalence. Many negative objects have positive features and vice versa; these features may set the stage for feature matching and source confusability. Any CS having both positive and negative features would seem to be a target for implicit misattribution of responses evoked by certain appropriate US (especially mildly evocative ones). The finding that the surveillance procedure was successful in reducing the negativity of automatically activated attitudes toward Black faces used as CS (Olson & Fazio, 2004) suggests that implicit misattribution can act on CS that are other than neutral. Another possibility is that preceding EC with attempts to weaken or de-automatize CS attitudes (Sanbonmatsu, Posavac, Vanous, Ho, & Fazio, 2007) may suffice to preempt some CS responses and open the door for source confusion. The viability of EC as an intervention for non-neutral CS is likely to depend heavily on the mechanism of EC, and each is likely to impose different boundary conditions on such effects. Interestingly, recent research in the domain of alcohol abuse has suggested that both a push-pull EC procedure (Wiers, Rinck, Kordts, Houben, & Strack, 2010) and a procedure similar to the surveillance procedure involving pairings of alcohol stimuli with valenced US pictures (Houben, Havermans, & Wiers, under review) influence alcohol attitudes, suggesting that intervention applications of EC are a promising avenue for future research.

VII. Concluding Thoughts

A. The Power of EC

The field has long sought to understand the origins of attitudes (Eagly & Chaiken, 1993), and contemporary theories of persuasion have provided key insights into how attitudes form through a variety of mechanisms (e.g., Chaiken, Lieberman, & Eagly, 1989; Petty & Cacioppo, 1981). However, these theories have focused most of their energies on conscious, deliberate, so-

called “explicit” processes, relegating less conscious and more implicit processes to the “periphery” (Briñol, Petty, & McCaslin, 2009). Such a focus is reasonable given evidence that persuasion via more thoughtful processes typically leads to stronger attitudes, and that persuasion brought about through less thoughtful processes tends to be ephemeral (e.g., Bizer & Krosnick, 2001; Petty, Haugtvedt, & Smith, 1995). Indeed, implicitly formed attitudes (like those created using our EC paradigm) are likely to be weak in at least one sense: they resemble a “gut” reaction without an underlying “reason.” Viewed in this way, such attitudes might be thought to be less resistant to counter-persuasion attempts, particularly those that evoke facts, figures, and the like. Thus, the conventional wisdom might suggest that attitudes produced through EC are not worthy of the attention of attitude researchers. Is there any reason to suspect, then, that implicitly formed attitudes like those developed though EC can have a reliable impact on judgments and behavior?

There is evidence suggesting that implicitly formed attitudes can be consequential. As described above, our earlier work demonstrated that attitudes produced through EC are capable of automatic activation (Olson & Fazio, 2002), one hallmark of attitude strength. Using a subliminal priming procedure, we demonstrated that attitudes toward the novel cartoon character CS were activated automatically upon their subliminal presentation, without any intentions on the part of respondents. A wealth of evidence highlights the power of attitudes capable of automatic activation (for a review see Fazio, 2007). Such attitudes can, for example, guide attention to the object (Roskos-Ewoldsen & Fazio, 1992), influence its construal (Smith, Fazio, & Cejka, 1996) and bias processing of object-related information (Houston & Fazio, 1989; Fazio, Ledbetter, & Towles-Schwen, 2000).

We have begun to accumulate evidence that the attitudes produced in our EC paradigm can have downstream information processing consequences like those described above. In one such study, participants were exposed to ambiguous attribute-related information about the CS after undergoing the usual conditioning procedure (Kendrick & Olson, 2009). The ambiguous information was presented in the form of Pokemon trading cards (the type of cartoon character that served as CS), which while perhaps decipherable to a Pokemon enthusiast, can be baffling to a novice. One card, for example, rated one of the CS as an "M" with reference to its "LRT," and indicated its "attack ability" as a "50HP." These ambiguous characteristics are meant to imply something about the object—its strengths, weaknesses, and other features. Participants were asked to disambiguate this information, i.e., to give their best guess as to what it meant, on a Likert-type scale. As expected, participants disambiguated the information in an attitude-congruent fashion. Across participants, the same information (e.g., 50 HP) was estimated to be positive if it described the CS they were conditioned to like, but it was seen as negative if it described the CS they were conditioned to dislike. This happened, of course, without participants' knowledge that their attitudes had been changed.

In a follow-up study, we extended these findings to attitude object-related memory. After conditioning, participants were provided with an extensive list of trait information about the CS, some positive, some negative, and some neutral. In a later recognition test, they were more likely to remember traits that corresponded to the direction of their conditioning: they correctly identified more positive information about the CS they were conditioned to like, and more negative information about the CS they were conditioned to dislike (Kendrick & Olson, 2009). Thus, attitudes formed through implicit EC appear to have biasing effects on later processing of

information about the attitude object, which, in turn, is likely to influence later judgments and behavior toward it.

Other researchers have also shown how impactful EC can be. For example, food preferences can be conditioned in such a way that influences later eating choices (Verhulst, Hermans, Baeyens, Spruyt, & Eelen, 2006). In other work, Custers and Aarts (2005) demonstrated that associating a goal with positive affect through EC causes people to want to achieve that goal, and to work more quickly through an intervening task that interferes with their ability to achieve it. Their EC procedure was entirely implicit, and employed different US (non-repeating words whose only commonality was their shared valence). Thus, evaluative conditioning can be applied to goals; associating a goal with positive affect—even implicitly—can cause people to pursue it. EC applications in marketing and advertising research have also demonstrated its potential to have noteworthy consequences for consumer psychology (e.g. Gibson, 2008; Strick, van Baaren, Holland, & Van Knippenberg, 2009). Finally, in an interesting application of the surveillance procedure, Nierman and Crandall (2010) paired relatively unfamiliar countries with positive and negative words and images to examine whether conditioned attitudes would lead to the generation of stereotypes about the citizens of those countries. They found that the US pairings influenced judgments of the nationality's warmth (but not competence), suggesting that EC can promote the development of specific stereotypic beliefs justifying the conditioned evaluative association.

In short, despite their second-class citizenship among major theories of persuasion, implicitly formed attitudes, like those produced through EC, can have real impact. By biasing the processing of subsequent information and choice behavior, such attitudes can come to be much more than a "gut" reaction without underlying "reason." The "gut" reactions can easily

come to be justified. Indeed, recent theorizing suggests that some deeply-rooted attitudes—attitudes that persist over time and resist counter persuasion—can have implicit origins (Rudman, 2004). The best single example of such attitudes, and their strength, may be irrational fears or phobias. Such negative attitudes can be very resistant to change, and even when interventions seem to succeed, relapse or "return of fear" is very common (Bouton, 2002; Rachman, 1989). Therefore, elucidating the mechanisms underlying attitude development, like what occurs in EC, is critical to providing a fuller answer to the question of how attitudes originate and potentially also how they might be changed.

B. Method and mechanism

A major goal of future research should be the furtherance of understanding of how specific methods relate to procedures. Designing an EC study requires numerous decisions about the choice of CS, US, number and parameters of pairings, use of a cover story, and so on. These decisions should increasingly be made in a manner driven by theory about the process underlying EC. For some purposes (e.g., some research uses EC because it is a useful way to create an attitude under experimental control), the underlying process may not be of primary interest. Nonetheless, we would note that because different mechanisms of EC produce attitudes with different sorts of bases, the operant mechanism may have practical consequences. For other purposes, that is, any research in which EC itself is the phenomenon of interest, it is critical to understand the process that is operating. Methodological decisions that aim to isolate a particular mechanism by rendering others unlikely to operate will further lead to theoretical clarity and empirical coherence.

Sweldens (2009) has developed an innovative approach to studying EC in which an experiment includes two variants of a procedure equated in as many ways as possible but

different in critical ways that dictate the likely mechanism underlying EC. Specifically, Sweldens has compared "sequential homogeneous" CS-US pairings to "simultaneous heterogeneous" pairings. Sequential homogeneous pairings involved each CS being followed after a delay by a single US. Simultaneous heterogeneous pairings involve a CS-US pairing having a shared onset and offset, and each CS is paired with various non-repeating US. Sweldens hypothesized that sequential homogeneous pairings would produce EC through a stimulus-stimulus association mechanism. He demonstrated that various characteristics of EC in these conditions did indicate that CS attitudes were being formed via such associations—for example, they were subject to US re-valuation (the phenomenon most clearly tied specifically to the CS-US association mechanism) and to retroactive interference, that is, new learning that makes newly formed associations relatively less accessible. These conditions also displayed dependence upon contingency awareness. The simultaneous heterogeneous conditions displayed very different characteristics. These are the same type of CS-US pairings that are used in the surveillance procedure, and the characteristics of EC in this set of conditions, which Sweldens described as involving "direct affect transfer," were consistent with implicit misattribution. EC was also successfully produced in these conditions, but neither US re-valuation nor retroactive interference appeared to operate. Further, EC operated in the absence of contingency awareness, which was carefully measured. This new approach suggests exciting ways to directly contrast EC operating through different mechanisms, and we hope to see more research that extends and refines it.

C. Implicit misattribution: Overview

In this paper, we have described a program of EC research that led to the formulation of the IMM, a novel approach to EC that can explain some of the properties that sometimes

characterize EC and distinguish it from PC. The IMM is innovative in that we would never have entertained manipulations like eye gaze shifting, evocativeness, and CS salience without such a theoretical perspective. We hope that the model and these findings, as well as our review of mechanisms of EC, will reduce the elusiveness of EC and clarify the source of earlier discrepancies in the literature. In our view, the EC literature has been unnecessarily contentious, and researchers from multiple perspectives have been too eager to dismiss carefully conducted research on conjecture when findings diverge from one's own perspective. We hope future theorizing can increasingly address *when* EC has particular characteristics rather than dealing in absolutes, especially for those who prefer to conceptualize EC as an effect. Alternatively, those taking the perspective that EC should be defined as a specific process should be explicit about that process and take care to avoid generalizing broadly beyond that process. In conclusion, there is considerable value to paying attention to the "how" question, that is, the examination of the processes that can mediate the effect of mere spatio-temporal co-occurrence with a US on attitude toward a CS. Multiple mechanisms can produce the effect, but which specific mechanism is playing the larger role in any given context is critical for related questions.

Author Note

Preparation of this chapter was supported by Grant MH38832 from the National Institute of Mental Health. The authors would like to thank Wil Cunningham and Eva Walther for comments on earlier drafts.

References

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice-Hall.

Ashby, F.G., Maddox, W.T. (2005). Human category learning. *Annual Review of Psychology, 56*, 149-178.

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., Hermans, R., & Eelen, P. (1993). The role of CS-US contingency in human evaluative conditioning. *Behavior Research and Therapy*, 31, 731-737.

Bem, D. J. (1972) Self-perception theory. In L. Berkowitz (ed), *Advances in Experimental Social Psychology*, (Vol. 6. pp. 1-62). New York: Academic Press.

Bizer, G. Y., & Krosnick, J. A. (2001). Exploring the structure of strength-related attitude features: The relation between attitude importance and attitude accessibility. *Journal of Personality and Social Psychology, 81*, 566-586.

Blechert, J., Michael, T., Williams, S. L., Purkis, H. M, & Wilhelm, F. H. (2008). When two paradigms meet: Does evaluative learning extinguish in differential fear conditioning? *Learning and Motivation, 39*, 58-70.

Bouton, M. E. (2002). Context, ambiguity, and unlearning: Sources of relapse after behavioral extinction. *Biological Psychiatry, 52*, 976-986.

Briñol, P., Petty, R. E., & McCaslin, M. (2009). Changing attitudes on implicit versus explicit measures: What is the difference? In R. E. Petty, R. H. Fazio, & P. Briñol (Eds.). *Attitudes: Insights from the new implicit measures* (pp. 285-326). New York, NY:

Psychology Press.

Cacioppo, J. T., Marshall-Goodell, B. S., Tassinary, L. G., & Petty, R. E. (1992). Rudimentary determinants of attitudes: Classical conditioning is more effective when prior knowledge about the attitude stimulus is low than high. *Journal of Experimental Social Psychology, 28,* 207-233.

Carlston, D. E. & Skowronski, J. J. (2005). Linking versus thinking: Evidence for the different associative and attributional bases of spontaneous trait transference and spontaneous trait inference. *Journal of Personality and Social Psychology, 89,* 884-898.

Cesario, J., Grant, H., & Higgins, E. T. (2004). Regulatory fit and persuasion: Transfer from "feeling right." *Journal of Personality and Social Psychology,* 388-404.

Chaiken, S., Liberman, A. & Eagly, A. H. (1989). Heuristic and systematic information processing within and beyond the persuasion context. In Uleman, J. S. & Bargh, J. A. (Eds.), *Unintended thought* (pp. 212-252). New York: Guilford Press.

Cohen, A., & Ivry, R. (1989). Illusory conjunctions inside and outside the focus of attention. *Journal of Experimental Psychology: Human Perception & Performance, 15,* 650-663.

Crawford, L. E., & Cacioppo, J. T. (2002). Learning where to look for danger: Integrating affective and spatial information. *Psychological Science,* 13, 449-453.

Croizet, J.-C., & Fiske, S. T. (2000). Moderation of priming by goals: Feeling entitled to judge increases the judged usability of evaluative primes. *Journal of Experimental Social Psychology, 36,* 155-181.

Custers, R., & Aarts, H. (2005). Positive affect as implicit motivator: On the nonconscious operation of behavioral goals. *Journal of Personality and Social Psychology, 89,* 129-142.

Davey, G. C. L. (1994). Defining the important questions to ask about evaluative conditioning: conditioning: A reply to Martin and Levey (1994). *Behavior Research and Therapy, 32,* 307-310.

Dawson, M. E., Rissling, A. J., Schell, A. M, & Wilcox, R. (2007). Under what conditions can human affective conditioning occur without contingency awareness? Test of the evaluative conditioning paradigm. *Emotion,* 755-766.

De Houwer, J. (Ed.). (2005). Evaluative conditioning [Special issue]. *Cognition and Emotion, 19*(*2*).

De Houwer, J. (2007). A conceptual and theoretical analysis of evaluative conditioning. *The Spanish Journal of Psychology, 10,* 230-241.

De Houwer, J. (2009). The propositional approach to associative learning as an alternative for association formation models. *Learning and Behavior, 37,* 1-20.

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

De Houwer, J., Baeyens, F., & Eelen, P. (1994). Verbal evaluative conditioning with undetected US presentations. *Behaviour Research and Therapy, 32,* 629-633.

De Houwer, J., Hendrickx, H., & Baeyens, F. (1997). Evaluative learning with "subliminally" presented stimuli. *Consciousness and Cognition, 6,* 87-107.

De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit measures: A normative analysis and review. *Psychological Bulletin, 135,* 347-368.

De Houwer, J., Thomas, S., & Baeyens, F. (2001). Associative learning of likes and dislikes: A review of 25 years of research on human evaluative conditioning. *Psychological Bulletin*, 127, 853-869.

Dijksterhuis, A. (2004). I like myself but I don't know why: Enhancing implicit self-esteem by subliminal evaluative conditioning. *Journal of Personality and Social Psychology, 86,* 345-355.

Dutton D. G. & Aron, A. P. (1974). Some evidence for heightened sexual attraction under conditions of high anxiety. *Journal of Personality and Social Psychology*, 30, 510-517.

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes.* Fort Worth, TX: Harcourt Brace Jovaovich.

Fazio, R. H. (2007). Attitudes as object-evaluation associations of varying strength. *Social Cognition, 25,* 603-637.

Fazio, R. H., Eiser, J. R., & Shook, N. J. (2004). Attitude formation through exploration: Valence asymmetries. *Journal of Personality and Social Psychology*, 87, 293-311.

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology*, 69, 1013-1027.

Fazio, R. H., Ledbetter, J. E., & Towles-Schwen, T. (2000). On the costs of accessible attitudes: Detecting that the attitude object has changed. *Journal of Personality and Social Psychology, 78,* 197-210.

Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology*, 54, 297-327.

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic

activation of attitudes. *Journal of Personality and Social Psychology*, 50, 229-238.

Field, A. P. (2000a). 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. (2000b). Evaluative conditioning is Pavlovian conditioning: Issues of definition, importance, and the theoretical importance of contingency awareness. *Consciousness and Cognition, 9,* 41-49.

Fishbein, M. (1963). An investigation of relationships between beliefs about an object and the attitude toward that object. *Human Relations*, 16, 233-240.

Fishbein, M. & Middlestadt, S. (1995). Noncognitive effects on attitude formation and change: Fact or artifact? *Journal of Consumer Psychology*, 4, 181-202.

Fulcher, E. P. & Hammerl, M. (2001). When all is revealed: A dissociation between evaluative learning and contingency awareness. *Consciousness and Cognition*, *10,* 524-549.

Gallant, J. L., & Garner, W. R. (1988). Some effects of distance and structure on conjunction errors. *Bulletin of the Psychonomic Society*, *26,* 323-326.

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

Gawronski, B. & LeBel, E. P. (2008). Understanding patterns of attitude change: When implicit measures show change but explicit measures do not. *Journal of Experimental Social Psychology*, *44*, 1355-1361.

Gibson, B. (2008). Can evaluative conditioning change attitudes toward mature brands? New evidence from the implicit association test. *Journal of Consumer Research*, *35*, 178-188.

Grumm, M., Nestler, S., & von Collani, G. (2009). Changing explicit and implicit attitudes: The case of self-esteem. *Journal of Experimental Social Psychology, 45*, 327-335.

Hammerl, M., Bloch, M., & Silverthorne, C. P. (1997). Effects of US-alone presentations on human evaluative conditioning. *Learning and Motivation, 28*, 491-509.

Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.

Hermans, D., Crombez, G., Vansteenwegen, D., Baeyens, F., & Eelen, P. (2002). Expectancy-learning and evaluative learning in human classical conditioning: Differential effects of extinction. In S. P. Shohov (Ed.), *Advances in Psychology Research* (Vol. 12, pp. 17-40), Hauppage, NY, US: Nova Science Publishers

Hermans, D., Vansteenwegen, D., Crombez, G., Baeyens, F., & Eelen, P. (2002). Expectancy learning and evaluative learning in human classical conditioning: Affective priming as an indirect and unobtrusive measure of conditioned stimulus valence. *Behavior Research and Therapy*, 40, 217-234.

Higgins, E. T. (2005). Value from regulatory fit. *Current Directions in Psychological Science, 14*, 209-213.

Hofmann, W., De Houwer, J., Perugini, M., Baeyens, F., & Crombez, G. (in press). Evaluative conditioning in humans: A meta-analysis. *Psychological Bulletin*.

Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis of the correlation between the implicit association test and explicit self-report measures. *Personality and Social Psychology Bulletin, 31*, 1369-1385.

Houben, K., Havermans, R. C., & Wiers, R. W. (under review). Learning to dislike alchohol: Conditioning negative implicit attitudes towards alcohol and its effect on drinking behavior.

Houston, D. A., & Fazio, R. H. (1989). Biased processing as a function `of attitude accessibility: Making objective judgments subjectively. *Social Cognition*, 7, 51-66.

Jefferis, V. E., & Fazio, R. H. (2008, February). Evaluative conditioning and avoidance behavior. Paper presented at the 2008 meeting of the Society for Personality and Social Psychology, Albuquerque, NM.

Jefferis, V. E., Loersch, C., & Fazio, R. H. (2007, January). Toward (and away from) the unknown: Evaluative conditioning of subliminally presented novel stimuli via motor behavior. Paper presented at the 2007 meeting of the Society for Personality and Social Psychology, Memphis, TN.

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,* 933-948.

Jones, E. E., & Nisbett, R. E. (1971). *The actor and the observer: Divergent perceptions of the causes of behavior*. New York: General Learning Press.

Katz, D. (1960). The functional study of attitudes. *Public Opinion Quarterly*, *24*, 163-204.

Kawakami, K., Phills, C. E., Steele, J. R., & Dovidio, J. F. (2007). (Close) distance makes the heart grow fonder: Improving implicit racial attitudes and interracial interactions through approach behaviors. *Journal of Personality and Social Psychology, 92,* 957-971.

Kendrick, R. V., & Olson, M. A. (2007, January). Following the heart versus the head in reporting implicitly formed attitudes. Paper presented at the 2007 meeting of the Society for Personality and Social Psychology, Memphis, TN.

Kendrick, R.V., & Olson, M.A. (2009, February). *The Downstream Effects of Implicitly-Formed Attitudes on Information Processing: When Gut Feelings Bias Perceptions.*

Paper presented at the 2009 meeting of the Society for Personality and Social Psychology, Tampa, FL.

Kim, J., Allen, C. T., & Kardes, F. R. (1996). An investigation of the mediational mechanisms underlying attitudinal conditioning. *Journal of Marketing Research, 33,* 318- 328.

Krosnick, J. A., Betz, A. L., Jussim, L. J., & Lynn, A. R. (1992). Subliminal conditioning of attitudes. *Personality and Social Psychology Bulletin, 18,* 152-162.

Kruschke, J. K, Kappenman, E. S., & Hetrick, W. P. (2005). Eye gaze and individual differences consistent with learned attention in associative blocking and highlighting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 830-845.

Lang, P.J., Bradley, M.M., & Cuthbert, B.N. (2008). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical Report A-8*. University of Florida, Gainesville, FL.

Levey, A. B., & Martin, I. (1975). Classical conditioning of human 'evaluative' responses. *Behaviour Research and Therapy*, 4, 205-207.

Lipp, O. V., Oughton, N., & LeLievre, J. (2003). Evaluative learning in Pavlovian conditioning: Extinct, but still there? *Learning and Motivation, 34,* 219-239.

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.

Malone, J. C. (2002). *Theories of learning: A historical approach.* Belmont, CA: Wadsworth.

Meersman, T., De Houwer, J., Baeyens, F., Randell, T., & Eelen P. (2005). Beyond evaluative conditioning? Searching for associative transfer of non-evaluative stimulus properties.

Cognition and Emotion, 19, 283-306.

Murphy, S. T., & Zajonc, R. B. (1993). Affect, cognition, and awareness: Affective priming with optimal and suboptimal stimulus exposures. *Journal of Personality and Social Psychology, 64,* 723- 739.

Nierman, A. & Crandall, C. (2010, January). *Stereotypes of warmth but not competence emerge as justifications of prejudice*. Paper presented at the 2010 meeting of the Society for Personality and Social Psychology, Las Vegas, NV.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84,* 231-259.

Olson, M. A. & Fazio, R. H. (2001). Implicit attitude formation through classical conditioning. *Psychological Science,* 12, 413-417.

Olson, M. A., & Fazio, R. H. (2002). Implicit acquisition and manifestation of classically conditioned attitudes. *Social Cognition, 20,* 89-103.

Olson, M. A., & Fazio, R. H. (2006). Reducing automatically-activated racial prejudice through implicit evaluative conditioning. *Personality and Social Psychology Bulletin, 32, 421-433.*

Olson, M. A., & Fazio, R. H. (2009). Implicit and explicit measures of attitudes: The perspective of the MODE model. In R. E. Petty, R. H. Fazio, & P. Briñol (Eds.), *Attitudes: Insights from the new implicit measures* (pp.19-63). New York, NY: Psychology Press.

Olson, M. A., Kendrick, R. V. & Fazio, R. H. (2009). Implicit covariation learning in evaluative vs. non-evaluative dimensions. *Journal of Experimental Social Psychology, 45,* 398-403.

Page, M. M. (1974). Demand characteristics and the classical conditioning of attitudes experiment. *Journal of Personality and Social Psychology*, 30, 468-476.

Payne, B. K., Burkley, M. A., & Stokes, M. B. (2008). Why do implicit and explicit attitude tests diverge? The role of structural fit. *Journal of Personality and Social Psychology, 94*, 16-31.

Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. (2005). An inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology, 89*, 277-293.

Petty, R. E., & Cacioppo, J. T. (1981). *Attitudes and persuasion: Classic and contemporary approaches.* Dubuque, IA: Wm. C. Brown.

Petty, R. E., Haugtvedt, C., & Smith, S. M. (1995). Elaboration as a determinant of attitude strength: Creating attitudes that are persistent, resistant, and predictive of behavior. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 93-130). Mahwah, NJ: Erlbaum.

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*, 130-144.

Pleyers, G., Corneille, O., Yzerbyt, V., & Luminet, O. (2009). Evaluative conditioning may incur attentional costs. *Journal of Experimental Psychology: Animal Behavior Processes, 35*, 279-285.

Powell, M. C., & Fazio, R. H. (1984). Attitude accessibility as a function of repeated attitudinal expression. *Personality and Social Psychology Bulletin*, 10, 139-148.

Pryor, J. B., & Kriss, M. (1977). The cognitive dynamics of salience in the attribution process. *Journal of Personality and Social Psychology, 35*, 49-55.

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

Rachman, S. J. (1989). The return of fear: Review and prospect. *Clinical Psychology Review, 9,* 147-168.

Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review,* 364-382.

Reber, R., Winkielman, P., & Schwarz, N. (1998). Effects of perceptual fluency on affective judgments. *Psychological Science, 9,* 45-48.

Rescorla, R. A. (1988). Pavlovian conditioning: It's not what you think it is. *American Psychologist, 43,* 151-160.

Roskos-Ewoldsen, D. R., & Fazio, R. H. (1992). On the orienting value of attitudes: Attitude accessibility as a determinant of an object's attraction of visual attention. *Journal of Personality and Social Psychology,* 63, 198-211.

Rozin, P. & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review, 5,* 296-320.

Rozin, P., Wrzesniewski, A., & Byrnes, D. (1998). The elusiveness of evaluative conditioning. *Learning and Motivation, 29,* 397-415.

Rudman, L. A. (2004). Sources of implicit attitudes. *Current Directions in Psychological Science, 13,* 80-83.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review, 110,* 145-172.

Ruys, K.I., Stapel, D.A. (2009). Learning to like or dislike by association: No need for

contingency awareness. *Journal of Experimental Social Psychology, 45*, 1277-1280.

Rydell, R. J. & Jones, C. R. (2009). Competition between unconditioned stimuli in attitude formation: Valence asymmetry versus spatio-temporal contiguity. *Social Cognition, 27*, 905-916.

Rydell, R. J., McConnell, A. R., Mackie, D., & Strain, L. (2006). Of two minds: Forming and changing valence-inconsistent implicit and explicit attitudes. *Psychological Science, 17*, 954-958.

Sanbonmatsu, D. M., Posavac, S. S., Vanous, S., Ho, E. A., & Fazio, R. H. (2007). The deautomatization of accessible attitudes. *Journal of Experimental Social Psychology, 43*, 365-378.

Schachter, S. & Singer, J. E. (1962). Cognitive, social, and psychological determinants of emotional state. *Psychological Review*, 69, 379-399.

Schwarz, N. & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology, 45*, 513-523.

Skonronski, J.J., Carlston, D. E., & Mae, L. (1998). Spontaneous trait transference: Communicators take on the qualities they describe in others. *Journal of Personality and Social Psychology, 74*, 837-848.

Smith, E. R., Fazio, R. H., & Cejka, M. A. (1996). Accessible attitudes influence categorization of multiply categorizable objects. *Journal of Personality and Social Psychology, 71*, 888-898.

Staats, A. W., & Staats, C. K. (1958). Attitudes established by classical conditioning. *Journal of Abnormal and Social Psychology, 57*, 37-40.

Stahl, C., Unkelbach, C. & Corneille, O. (2009). On the respective contributions of awareness of US valence and US identity in attitude formation through evaluative conditioning. *Journal of Personality and Social Psychology, 97,* 404-420

Storms, M. D. (1973). Videotape and the attribution process: Reversing actors' and observers' points of view. *Journal of Personality and Social Psychology*, *27,* 165-175.

Storms, M. D., & Nisbett, R. E. (1970). Insomnia and the attribution process. *Journal of Personality and Social Psychology, 16,* 319-328.

Strack, F. & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, *8*, 220-247.

Strick, M., van Baaren, R. B., Holland, R. W., & van Knippenberg, A. (2009). Humor in advertisements enhances product liking by mere association. *Journal of Experimental Psychology: Applied*, *15,* 35-45.

Sweldens, S. (2009*). Evaluative conditioning 2.0: Direct versus associative transfer of affect to brands*. Unpublished doctoral dissertation, Erasmus University, NL.

Taylor, S. E., & Fiske, S. T. (1978). Salience, attention, and attribution: Top-of-the-head phenomena. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 11). New York: Academic Press.

Thompson, W. F., Hall, M. D., & Pressing, J. (2001). Illusory conjunctions of pitch and duration in unfamiliar tone sequences. *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 128-140.

Till, B. D., Stanley, S. M., & Priluck, R. (2008). Classical conditioning and celebrity endorsers: An examination of belongingness and resistance to extinction. *Psychology & Marketing*, *25*, 179-196.

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

Treisman, A. M. & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology, 12,* 97-136.

Treisman, A. M. & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology, 14,* 107-141.

Vaughn, L. A., Malik, J., Schwartz, S., Petkova, Z., & Trudeau, L. (2006). Regulatory fit as input for stop rules. *Journal of Personality and Social Psychology, 91,* 601-611.

Verhulst, F., Hermans, D., Baeyens, F., Spruyt, A., & Eelen, P. (2006). Determinants and predictive validity of direct and indirect measures of recently acquired food attitudes. *Appetite, 46,* 137-43.

Walther, E. (2002). Guilty by mere association: Evaluative conditioning and the spreading attitude effect. *Journal of Personality and Social Psychology, 82,* 919-934.

Walther, E., Gawronski, B., Blank, H., & Langer, T. (2009). Changing likes and dislikes through the back door: The US-revaluation effect. *Cognition and Emotion, 23,* 889-917.

Walther, E., & Grigoriadis, S. (2004). Why sad people like shoes better: The influence of mood on the evaluative conditioning of consumer attitudes. *Psychology & Marketing, 21,* 755-773.

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,* 454-459.

Wegener, D. T., & Petty, R. E. (1995). Flexible correction processes in social judgment: The

role of naïve theories in corrections for perceived bias. *Journal of Personality and Social Psychology*, *68*, 36-51.

Weiner, B. (1985). "Spontaneous" Causal Thinking. *Psychological Bulletin*, *97*, 74-84.

Whitfield, M. & Jordan, C. H. (in press). Mutual influence of implicit and explicit attitudes. *Journal of Experimental Social Psychology*.

Wiers, R. W., Rinck, M., Kordts, R., Houben, K., & Strack, F. (2010). Retraining automatic action-tendencies to approach alcohol in hazardous drinkers. *Addiction*, doi:10.1111/j.1360-0443.2009.02775.x

Wilson, T. D. & Brekke, N. (1994). Mental contamination and mental correction: Unwanted influences on judgments and evaluations. *Psychological Bulletin*, *116*, 117-142.

Winkielman, P. & Cacioppo, J. T. (2001). Mind at ease puts a smile on the fade: Psychophysiological evidence that processing facilitation elicits positive affect. *Journal of Personality and Social Psychology*, *81*, 989-1000.

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, *35*, 151-175.

Zanna, M. P. & Cooper, J. (1974). Dissonance and the pill: An attribution approach to studying the arousal properties of dissonance. *Journal of Personality and Social Psychology*, *29*, 703-709.

Zanna, M. P., Kiesler, C. A., & Pilkonis, P. A. (1970). Positive and negative attitudinal affect established by classical conditioning. *Journal of Personality and Social Psychology*, 14, 321-328.

Zillman, D., Katcher, A. H., & Milavsky, B. (1972). Excitation transfer from physical exercise to subsequent aggressive behavior. *Journal of Experimental Social Psychology*, *8*, 247-259.