Article

Social Psychological and Personality Science
2017, Vol. 8(8) 858-866

DOI: 10.1177/1948550617691094
journals.sagepub.com/home/spp
SAGE

Propositional Versus Dual-Process Accounts of Evaluative Conditioning: II. The Effectiveness of Counter-Conditioning and Counter-Instructions in Changing Implicit and Explicit Evaluations

Xiaoqing Hu[1,2], Bertram Gawronski[3], and Robert Balas[4]

Abstract

Evaluative conditioning (EC) is defined as the change in the evaluation of a conditioned stimulus (CS) due to its pairing with a valenced unconditioned stimulus (US). Expanding on the debate between dual-process and propositional accounts, two studies investigated the relative effectiveness of counter-conditioning and counter-instructions in reversing EC effects on implicit and explicit evaluations. After conditioned evaluations were acquired via CS-US pairings, participants were either (1) presented with repeated CS-US pairings of the opposite valence or (2) given instructions that the CSs will be paired with USs of the opposite valence. Although both procedures reversed previously conditioned explicit evaluations, only directly experienced CS-US pairings reversed previously conditioned implicit evaluations. The findings question the functional equivalence of counter-conditioning and counter-instructions hypothesized by single-process propositional accounts. Yet, they support dual-process accounts, suggesting that associative and propositional processes jointly contribute to EC effects.

Keywords

associative learning, attitude change, evaluative conditioning, implicit evaluation, propositional learning

Uncovering the mechanisms underlying attitude formation and change is arguably one of the most important missions in understanding the human mind and behavior. Our attitudes toward objects and other people can guide a wide range of important decisions, including consumer behavior, food choice, and voting decisions. At an interpersonal level, preferences play a similarly critical role in starting a friendship, choosing a romantic partner, and interacting with out-group members. Because attitudes are involved in almost every aspect of social life, a substantial amount of research has been devoted to investigating how people's attitudes can be formed and changed.

Among the various phenomena in the literature on attitude formation and change, one particularly influential phenomenon is known as evaluative conditioning (EC). According to De Houwer (2007), EC can be defined as the change in the evaluation of a conditioned stimulus (CS) due to its pairing with a valenced unconditioned stimulus (US). Different from earlier conceptualizations in terms of associative link formation (see De Houwer, Thomas, & Baeyens, 2001), De Houwer's definition conceptualizes EC as a behavioral effect (i.e., the effect of stimulus pairings on evaluations), treating the mental processes underlying EC as a theoretical issue that has to be addressed on the basis of empirical data.

Indeed, recent research on EC has been shaped by intense debates regarding its underlying mental mechanisms (e.g., De Houwer, 2009; Gast, Gawronski, & De Houwer, 2012; Gawronski & Bodenhausen, 2011; Jones, Olson, & Fazio, 2010). One particularly controversial question in this debate concerns the role of associative processes in EC. Rejecting the idea of automatic association formation, single-process propositional accounts posit that EC effects are mediated by the non-automatic formation and truth assessment of mental propositions about the relation between the CS and the US (e.g., De Houwer, 2009, 2014; Mitchell, De Houwer, & Lovibond, 2009). In contrast, dual-process accounts assume that, in addition to propositional processes, EC effects can result from

[1] Department of Psychology, The University of Hong Kong, Hong Kong, China
[2] The State Key Laboratory of Brain and Cognitive Science, The University of Hong Kong, Hong Kong, China
[3] The University of Texas at Austin, Austin, TX, USA
[4] Polish Academy of Sciences, Warsaw, Poland

Corresponding Author:
Xiaoqing Hu, Department of Psychology, The University of Hong Kong, Pokfulam Road, Hong Kong.
Email: xiaoqinghu@hku.hk

the automatic formation of associative links between co-occurring stimuli (e.g., Gawronski & Bodenhausen, 2011; Sweldens, Van Osselaer, & Janiszewski, 2010).

Propositional Versus Dual-Process Accounts of EC

The single-process propositional model by De Houwer and colleagues (De Houwer, 2009, 2014; Mitchell et al., 2009) proposes that all instances of EC are mediated by the nonautomatic formation and truth assessment of mental propositions about the relation between the CS and the US. A central implication of this assumption is that the basis of these propositions is irrelevant for their effect on evaluative responses to the CS. From a propositional view, it does not matter whether a mental proposition has its origin in actually observed CS-US pairings or verbal information about CS-US pairings (De Houwer, 2009). Put differently, the effects of actually experienced CS-US pairings and verbal information about CS-US pairings on evaluative responses are assumed to be functionally equivalent, in that both effects are mediated by mental propositions about CS-US relations.

In contrast, dual-process accounts assume that EC effects can be the result of either associative or propositional processes, or both. For example, Gawronski and Bodenhausen's (2011) associative-propositional evaluation (APE) model acknowledges the hypothesized role of propositional processes. Yet, the theory further assumes that CS-US pairings can lead to EC effects via the automatic formation of associative links. According to the APE model, one of the key differences between associative and propositional processes is their relative effectiveness in overriding the effect of previously observed CS-US pairings on different kinds of evaluative responses. Although propositional inferences may lead to a rejection of previously formed associations for deliberate evaluative judgments, such a rejection is assumed to be ineffective in qualifying the effect of previously formed associations on spontaneous evaluative responses (Gawronski & Bodenhausen, 2006). Changes in spontaneous evaluative responses are assumed to require the formation of new associations on the basis of repeated stimulus co-occurrences.

Counter-Conditioning Versus Counter-Instructions

Propositional and dual-process accounts lead to conflicting predictions about the relative effectiveness of counter-conditioning (i.e., exposure to repeated CS-US pairings of the opposite valence) and counter-instructions (i.e., verbal information that the CSs will be paired with USs of the opposite valence without direct exposure to CS-US pairings) in reversing EC effects of previously observed CS-US pairings on spontaneous evaluative responses (i.e., implicit evaluations) and deliberate evaluative judgments (i.e., explicit evaluations). According to single-process propositional accounts, effects of counter-conditioning and counter-instructions should be functionally equivalent, in that both are mediated by mental propositions about CS-US relations. Hence, to the extent that either procedure is effective in influencing mental propositions, counter-conditioning and counter-instructions should be equally effective in reversing EC effects on implicit and explicit evaluations. There should be no differences for the two types of reversal procedure and the two types of evaluative responses.

In contrast, dual-process accounts such as the APE model suggest that the type of evaluative response is essential for understanding the relative effectiveness of counter-conditioning and counter-instructions (Gawronski & Bodenhausen, 2011). Although actually experienced CS-US pairings and verbal information about CS-US pairings may be equally effective in reversing EC effects on explicit evaluations, only actually experienced CS-US pairings, but not verbal instructions about CS-US pairings, should reverse EC effects on implicit evaluations. According to dual-process accounts, changes in the latter cannot be achieved by a rejection of previously formed associations on the basis of verbal information but require the formation of new associations as a result of repeated stimulus co-occurrences.

To clarify the difference between the two accounts, it is important to note that they do not lead to conflicting predictions about the effectiveness of actually experienced CS-US pairings and verbal instructions about CS-US pairings in the absence of previously observed CS-US pairings (e.g., De Houwer, 2006; Gast & De Houwer, 2012). According to the APE model, propositional inferences can influence implicit evaluations when they do not have to override effects of preexisting associations (see Gawronski & Bodenhausen, 2006, Case 4). In this case, actually experienced CS-US pairings and verbal instructions about CS-US pairings are assumed to be equally effective in influencing implicit evaluations. The critical difference between the two accounts concerns cases in which verbal information about CS-US pairings conflicts with previously observed CS-US pairings (cf. Gawronski & LeBel, 2008). In this case, dual-process accounts assume that the effect of previously formed associations on implicit evaluations has to be counteracted by the formation of new associations, which requires exposure to repeated stimulus pairings. Verbal information alone is insufficient to override the effect of preexisting associations on implicit evaluations (see Gawronski & Bodenhausen, 2006, Case 3). Thus, although newly acquired verbal information may qualify EC effects on explicit evaluations, the impact of previously formed associations on implicit evaluations should be unaffected.

The Present Research

The current research addressed this question by directly comparing the relative effectiveness counter-conditioning (i.e., exposure to repeated CS-US pairings of the opposite valence) and counter-instructions (i.e., verbal information that the CSs will be paired with USs of the opposite valence without direct exposure to CS-US pairings) in reversing EC effects of previously observed CS-US pairings. According to single-process propositional accounts, counter-conditioning and counter-instructions should be equally effective in reversing EC effects on implicit evaluations. In contrast, dual-process accounts suggest that only actually experienced CS-US pairings, but not verbal instructions about CS-US pairings, should reverse EC effects implicit evaluations. The two accounts lead to the same prediction for explicit evaluations, in that EC effects on explicit evaluations should be reversed by both counter-conditioning and counter-instructions.

To test these predictions, participants were presented with CS-US pairings and then completed measures of implicit and explicit evaluations of the CSs. After completion of the two evaluation measures, participants in the counter-conditioning group were presented with CS-US pairings of the opposite valence. Participants in the counter-instructions group were told that the CSs would be presented with USs of the opposite valence. Finally, all participants completed the two evaluation measures a second time.[1]

Experiment 1

Method

Participants and Design

One hundred and one psychology undergraduates were recruited for a 1 hr battery entitled "impression formation and moral judgments" that included the current experiment and two unrelated studies. Participants received research credit for an introductory psychology course. Due to a computer malfunction, data from one participant were lost. Four additional participants were excluded due to excessive invalid trials (>60%) in one of the implicit evaluation measures. This left us with a final sample of 96 participants (62 women) for the current analyses. The study included a 2 (initial US valence: positive vs. negative, within-subjects) $\times$ 2 (time of measurement: Time 1 vs. Time 2, within-subjects) $\times$ 2 (reversal procedure: counter-conditioning vs. counter-instructions, between-subjects) mixed design.

Materials

As CSs, we adapted 10 computer-generated images of shapes with different color patterns from Gawronski, Balas, and Creighton (2014). As USs, we used eight positive and eight negative pictures from the International Affective Picture System (Lang, Bradley, & Cuthbert, 2008).

EC Procedure

Four of the CSs were paired with one of the positive pictures as the US; four of the CSs were paired with one of the negative pictures as the US; and two of the CSs were not paired with a valenced picture to serve as neutral baseline stimuli in the evaluation measures (see below). The EC procedure included 10 presentations of each CS-US pair, summing up to a total of 80 trials. The stimuli of each CS-US pair were presented simultaneously for 1,000 ms, with the CSs being presented slightly below and the USs slightly above the center of the screen. The intertrial interval was 2,000 ms. The particular pairings of CSs and USs were counterbalanced with a Latin square. The verbatim instructions are available in the Supplemental Online Materials (SOMs).

Measures

To measure explicit evaluations, participants were shown each of the 10 CSs and asked to respond to the question "How pleasant or unpleasant do you find this image?" on a 7-point scale ranging from 1 (*very unpleasant*) to 7 (*very pleasant*). To measure implicit evaluations, we employed an evaluative priming task (Fazio, Jackson, Dunton, & Williams, 1995) that included the 10 CSs as primes and 10 positive and 10 negative adjectives as targets (for target words, see SOMs). Each trial started with a fixation cross that was displayed for 500 ms in the center of the screen. The fixation cross was followed by a prime, which was replaced by the target word after 200 ms. Participants' task was to press a right-hand key (*Numpad 5*) as quickly as possible when the target word was positive and a left-hand key (*A*) when the target word was negative. The target words remained on the screen until participants made their response. Incorrect responses were followed by the word *ERROR*! for 1,500 ms before the next trial started. The intertrial interval was 500 ms. Each of the 10 primes was presented once with each of the 10 positive target words and once with each of the 10 negative words, summing up to a total of 200 trials.

Counter-Conditioning

The counter-conditioning procedure was identical to the EC procedure, the only difference being that CSs that had been paired with a positive US were now paired with a negative US, and vice versa. The verbatim instructions are available in the SOMs.

Counter-Instructions

Different from the counter-conditioning procedure, participants in the counter-instructions group were not presented with any CS-US pairings. Instead, participants were presented with each CS and told whether it would be paired with a positive or a negative image (e.g., "This drawing will be presented with a POSITIVE image"). The verbatim instructions are available in the SOMs (cf. Gast & De Houwer, 2013). The instructed valence was always opposite to the valence of the US a given CS had been paired with in the initial EC procedure. Participants were not presented with any US images here.

Procedure

Participants were presented with CS-US pairings and then completed the measures of implicit and explicit CS evaluations in counterbalanced order. Afterward, half of the participants were presented with CS-US pairings of the opposite valence; the remaining half was told that the CSs would be presented

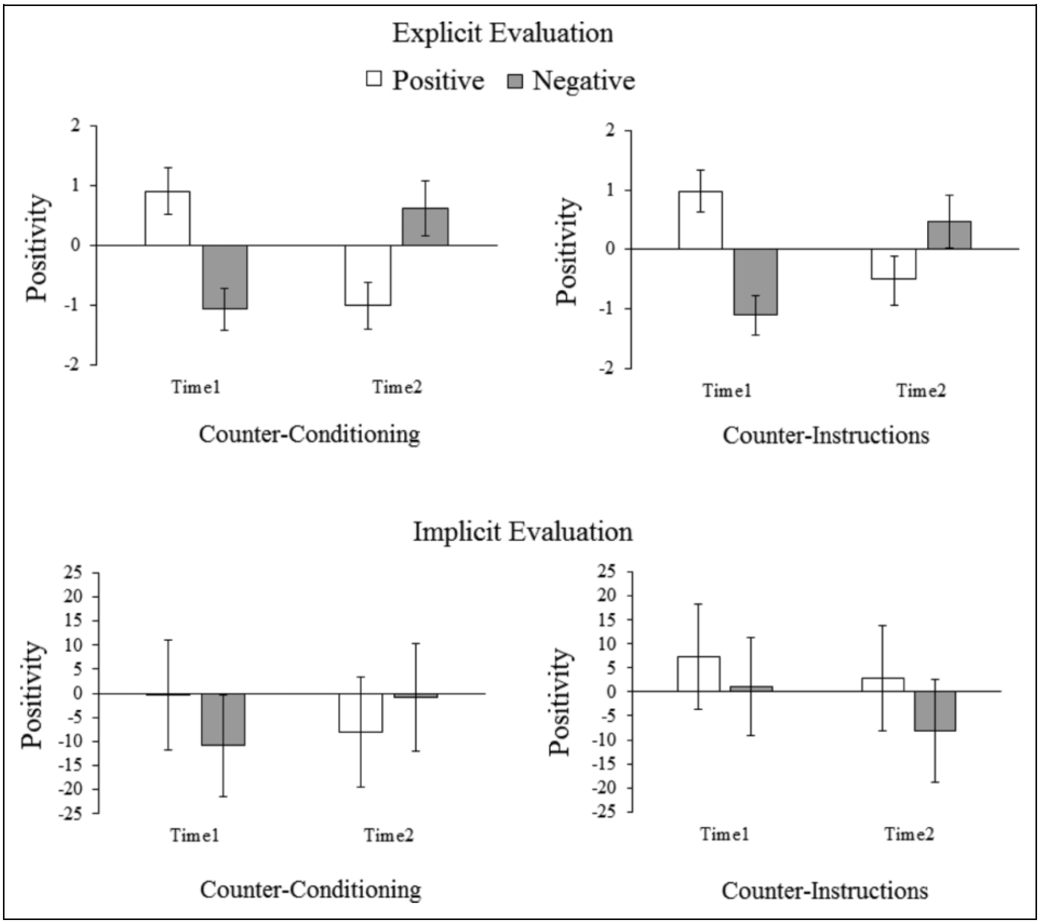

Figure 1. Explicit and implicit conditioned stimulus evaluations at Times 1 and 2 as a function of initial unconditioned stimulus valence (positive vs. negative) and reversal procedure (counter-conditioning vs. counter-instructions), Experiment 1. Error bars indicate 95% confidence intervals.

with a US of the opposite valence. Finally, participants completed the two evaluation measures a second time in the same order as in the first measurement.

Results

Data Aggregation

Baseline-corrected scores of explicit CS evaluations were calculated by averaging participants' ratings of all CSs that had been paired with a US of the same valence and then subtracting the mean ratings of the two baseline CSs from the two average scores. Thus, higher values indicate more favorable explicit evaluations of a given CS type compared to baseline. To obtain baseline-corrected scores of implicit CS evaluations, we first excluded trials with incorrect responses (4.8%) and truncated latencies higher than 800 ms (12.3%; see Hu, Gawronski, & Balas, 2017). We then calculated a positivity index for each CS type by subtracting the mean response latency to positive target words preceded by a given CS type from the mean response latency to positive target words preceded by the neutral baseline CSs (Wentura & Degner, 2010). Negativity indices were calculated accordingly by subtracting the mean response latency to negative target words preceded by a given CS type from the mean response latency to negative target words preceded by the neutral baseline CSs. Negativity scores were then subtracted from positivity scores for each of the two CS types. Thus, higher values indicate more favorable implicit evaluations of a given CS type compared to baseline.

Explicit Evaluations

Submitted to a 2 (initial US valence) × 2 (time of measurement) × 2 (reversal procedure) mixed-model analysis of variance (ANOVA), explicit evaluations revealed a significant main effect of Initial US Valence, $F(1, 94) = 10.70$, $p = .001$, $\eta_p^2 = .102$, which was qualified by a significant two-way interaction of Initial US Valence and Time of Measurement, $F(1, 94) = 106.62$, $p < .001$, $\eta_p^2 = .531$ (see Figure 1). Further analyses specified this interaction by revealing a regular EC effect at Time 1, $t(95) = 11.92$, $p < .001$, $d = 1.21$, and a reversed EC effect at Time 2, $t(95) = -5.86$, $p < .001$, $d = -.60$. The three-way interaction of Initial US Valence, Time of Measurement, and Reversal Procedure was not significant, $F(1, 94) = 0.71$, $p = .40$, $\eta_p^2 = .007$. The reversal of EC effects was statistically significant for both counter-conditioning and counter-instructions, as indicated by significant two-way interactions of Initial US Valence and Time of Measurement in the counter-conditioning group, $F(1, 45) = 56.27$, $p < .001$, $\eta_p^2 = .556$, and the counter-instructions group, $F(1, 49) = 49.85$, $p < .001$, $\eta_p^2 = .504$.

Implicit Evaluations

The same ANOVA on implicit evaluations yielded a marginally significant main effect of Initial US Valence, $F(1, 94) = 3.55$, $p = .06$, $\eta_p^2 = .036$, which was qualified by a significant three-way interaction of Initial US Valence, Time of Measurement, and Reversal Procedure, $F(1, 94) = 6.52$, $p = .01$, $\eta_p^2 = .065$ (see Figure 1). To specify this interaction, we conducted separate 2 (initial US valence) × 2 (time of measurement) repeated-measures ANOVAs for the counter-conditioning and counter-instructions groups, respectively.

For the counter-conditioning group, the ANOVA revealed a significant two-way interaction of Initial US Valence and Time of Measurement, $F(1, 45) = 9.79$, $p = .003$, $\eta_p^2 = .179$, indicating a marginal EC effect at Time 1, $t(45) = 1.95$, $p = .06$, $d = .29$, and a marginal reversed EC effect at Time 2, $t(45) = -1.92$, $p = .06$, $d = -.28$. In contrast, for the counter-instructions group, the ANOVA revealed only a significant main effect of Initial US Valence, $F(1, 49) = 4.80$, $p = .03$, $\eta_p^2 = .089$, which was unqualified by Time of Measurement, $F(1, 49) = 0.52$, $p = .47$, $\eta_p^2 = .011$.

Broken down by Time of Measurement, a 2 (initial US valence) × 2 (reversal procedure) mixed-model ANOVA on implicit evaluations at Time 1 revealed a significant main effect of Initial US Valence, $F(1, 94) = 4.88$, $p = .03$, $\eta_p^2 = .049$, which remained unqualified by Reversal Procedure, $F(1, 94) = 0.32$, $p = .57$, $\eta_p^2 = .003$. Critically, the same ANOVA on implicit evaluations at Time 2 revealed a significant two-way interaction, $F(1, 94) = 8.23$, $p = .005$, $\eta_p^2 = .080$, indicating a regular EC effect in the counter-instructions group, $t(49) = 2.19$, $p = .03$, $d = .31$, and a marginal reversed EC effect in the counter-conditioning group, $t(45) = -1.92$, $p = .06$, $d = -.28$.

Discussion

Experiment 1 obtained preliminary evidence that counter-conditioning and counter-instructions are differentially effective in reversing EC effects of previously observed CS-US pairings. Although both procedures reversed EC effects on explicit evaluations, only counter-conditioning was effective in reversing EC effects on implicit evaluations. These results conflict with the predictions of propositional accounts, which suggest that counter-conditioning and counter-instructions should be equally effective in reversing EC effects on implicit evaluations. Yet, the current findings are consistent with dual-process accounts, which suggest that a reversal of EC effects on implicit evaluations cannot be achieved by counter-instructions but instead requires repeated exposure to CS-US pairings of the opposite valence.

Although Experiment 1 supports dual-process accounts of EC, it seems desirable to replicate the main findings and explore their generality across different EC procedures. Recent evidence suggests that associative processes might play a less dominant role in EC effects resulting from sequential CS-US pairings as compared to EC effects resulting from simultaneous CS-US pairings (e.g., Hütter & Sweldens, 2013; Sweldens et al., 2010). Thus, counter-instructions could be more effective in reversing EC effects of sequential pairings. Experiment 2 addressed this question.

Experiment 2

Method

Participants and Procedures

One hundred and twenty-four undergraduates participated in the study for research credit. Due to computer malfunction, data from two participants were lost. Three additional participants were excluded due to excessive invalid trials (>60%) in one or both of the evaluative priming tasks. This left us with a final sample of 119 participants (74 women). The design, measures, and materials were identical to the ones in Experiment 1, the only difference being that the EC procedure and the counter-conditioning procedure used sequential CS-US pairings instead of simultaneous CS-US pairings. On each trial, a CS was presented for 1,000 ms, followed by a US for 1,000 ms. Both stimuli were presented in the center of the screen. The intertrial interval was 2,000 ms.

Results

Data Aggregation

Data were aggregated using the procedures of Experiment 1.

Explicit Evaluations

Submitted to a 2 (initial US valence) × 2 (time of measurement) × 2 (reversal procedure) mixed-model ANOVA, explicit evaluations revealed a significant main effect of Initial US Valence, $F(1, 117) = 7.01$, $p = .009$, $\eta_p^2 = .057$, which was qualified by a significant two-way interaction of Initial US Valence and Time of Measurement, $F(1, 117) = 152.35$, $p < .001$, $\eta_p^2 = .566$ (see Figure 2). Replicating Experiment 1, further analyses revealed a regular EC effect at Time 1, $t(118) = 13.06$, $p < .001$, $d = 1.20$, and a reversed EC effect at Time 2, $t(118) = -7.58$, $p < .001$, $d = -.69$. The three-way interaction of Initial US Valence, Time of Measurement, and Reversal Procedure was not significant, $F(1, 117) = 0.13$, $p = .722$, $\eta_p^2 = .001$. The reversal of EC effects was statistically significant for both counter-conditioning and counter-instructions, as indicated by significant two-way interactions

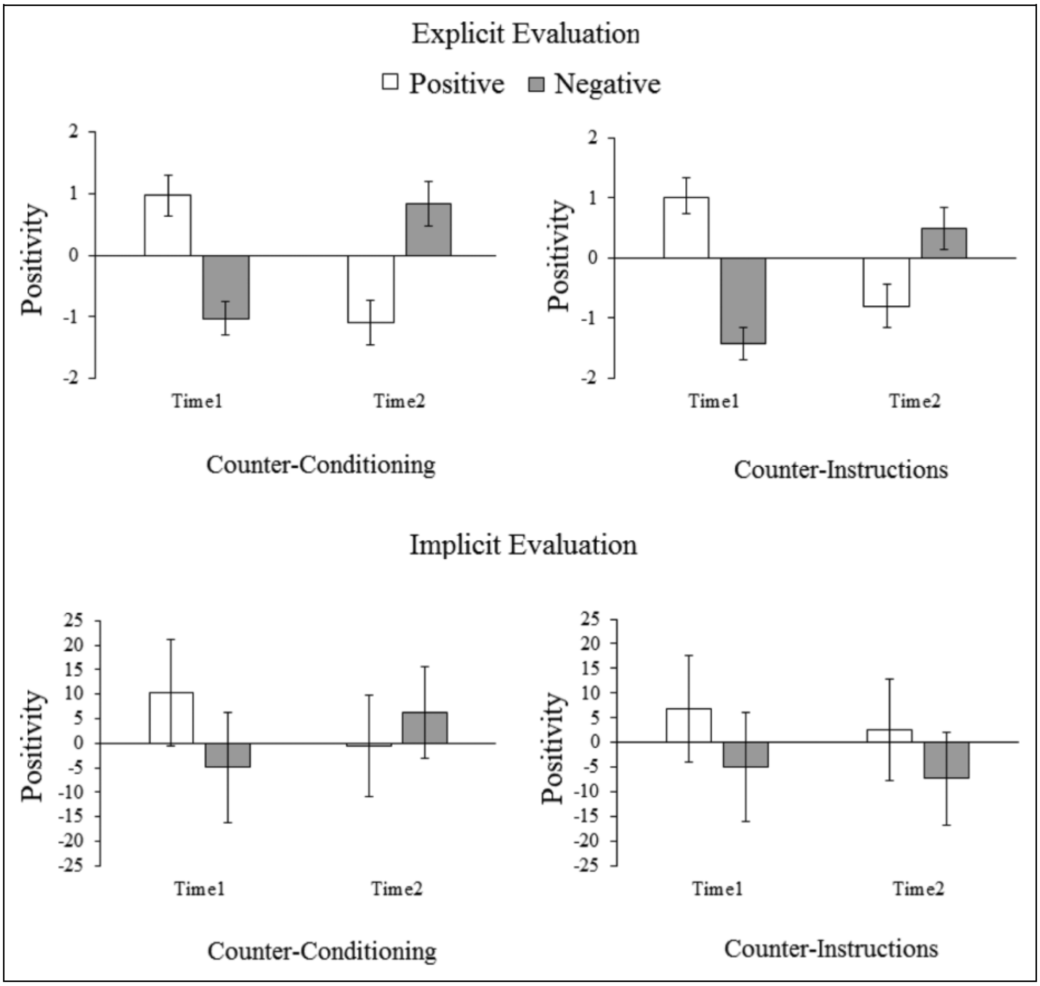

Figure 2. Explicit and implicit conditioned stimulus evaluations at Times 1 and 2 as a function of initial unconditioned stimulus valence (positive vs. negative) and reversal procedure (counter-conditioning vs. counter-instructions), Experiment 2. Error bars indicate 95% confidence intervals.

of Initial US Valence and Time of Measurement in the counter-conditioning group, $F(1, 58) = 69.54$, $p < .001$, $\eta_p^2 = .545$, and the counter-instructions group, $F(1, 59) = 84.97$, $p < .001$, $\eta_p^2 = .590$.

Implicit Evaluations

The same ANOVA on implicit evaluations yielded a significant main effect of Initial US Valence, $F(1, 117) = 9.03$, $p = .003$, $\eta_p^2 = .072$, and a significant two-way interaction of Initial US Valence and Time of Measurement, $F(1, 117) = 8.86$, $p = .004$, $\eta_p^2 = .070$, which were qualified by a significant three-way interaction of Initial US Valence, Time of Measurement, and Reversal Procedure, $F(1, 117) = 6.29$, $p = .01$, $\eta_p^2 = .051$ (see Figure 2). To specify this interaction, we conducted separate 2 (initial US valence) $\times$ 2 (time of measurement) repeated-measures ANOVAs for the counter-conditioning and counter-instructions groups, respectively.

For the counter-conditioning group, the ANOVA revealed a significant two-way interaction of Initial US Valence and Time of Measurement, $F(1, 58) = 13.89$, $p < .001$, $\eta_p^2 = .193$, indicating a significant EC effect at Time 1, $t(58) = 3.57$, $p = .001$, $d = .46$, and a marginal reversed EC effect at Time 2, $t(58) = -1.72$, $p = .09$, $d = -.22$. In contrast, for the counter-instructions group, the ANOVA revealed only a significant main effect of Initial US Valence, $F(1, 59) = 6.79$, $p = .01$, $\eta_p^2 = .103$, which was unqualified by Time of Measurement, $F(1, 59) = 0.12$, $p = .73$, $\eta_p^2 = .002$.

Broken down by Time of Measurement, a 2 (initial US valence) $\times$ 2 (reversal procedure) mixed-model ANOVA on implicit evaluations at Time 1 revealed a significant main effect of Initial US Valence, $F(1, 117) = 19.36$, $p < .001$, $\eta_p^2 = .142$, which remained unqualified by Reversal Procedure, $F(1, 117) = 0.33$, $p = .57$, $\eta_p^2 = .003$. Critically, the same ANOVA on implicit evaluations at Time 2 revealed a significant two-way interaction of Initial US Valence and Reversal Procedure, $F(1, 117) = 6.08$, $p = .02$, $\eta_p^2 = .049$, indicating a marginal EC effect for the counter-instructions group, $t(59) = 1.81$, $p = .08$, $d = .23$, and a marginal reversed EC effect in the counter-conditioning group, $t(58) = -1.72$, $p = .09$, $d = -.22$.

Discussion

To test the generality of the effects obtained in Experiment 1, Experiment 2 investigated the effectiveness of counter-conditioning and counter-instructions in reversing EC effects of sequential CS-US pairings. Replicating the findings obtained from simultaneous pairings, both counter-conditioning and counter-instructions reversed EC effects on explicit evaluations. However, only counter-conditioning, but not counter-instructions, was effective in reversing EC effects on implicit evaluations.

General Discussion

The current experiments suggest that counter-conditioning and counter-instructions are differentially effective in reversing EC effects of previously observed CS-US pairings. Although both procedures reversed EC effects on explicit evaluations, EC effects on implicit evaluations were reversed only by counter-conditioning but not counter-instructions. These results are consistent with dual-process accounts of EC, which suggest that a reversal of EC effects on implicit evaluations requires the formation of new associations as a result of repeated stimulus co-occurrences. However, they are inconsistent with propositional accounts, which suggest that counter-conditioning and counter-instructions are functionally equivalent.

Because some of the obtained effects were only marginal, we conducted a combined analysis of the two experiments to increase statistical power (see SOMs). The results of this analysis corroborate our conclusions, showing that EC effects on explicit evaluations were reversed by both counter-conditioning and counter-instructions. In contrast, EC effects on implicit evaluations were significantly reversed by counter-conditioning but not by counter-instructions.

Our findings expand on earlier research by Gast and De Houwer (2013) who investigated the effectiveness of counter-instructions in reversing the effects of initial instructions about CS-US pairings and actually observed CS-US pairings. In their research, EC effects on implicit evaluations were reduced as a result of counter-instructions. However, counter-instructions did not lead to a reversal and initial EC effects remained statistically significant after counter-instructions. Emphasizing the obtained effects of mere instructions, Gast and De Houwer interpreted their findings as evidence for propositional accounts of EC. However, their study did not compare the relative effectiveness of counter-instructions and counter-conditioning, which is essential in the debate between dual-process and propositional accounts. Dual-process accounts predict that (a) verbal information about CS-US pairings should reverse EC effects on explicit, but not implicit, evaluations, and (b) actually experienced CS-US pairings should reverse EC effects on both explicit and implicit evaluations. In contrast, propositional accounts predict that counter-conditioning and counter-instructions should be equally effective in reversing EC effects on explicit and implicit evaluations. The current findings support the prediction of dual-process accounts, but they are inconsistent with the predictions of single-process propositional accounts.

Potential Objections

A potential objection is that counter-instructions in the current studies were presented only once for each CS. Thus, their effects might have been weaker compared to the effects of counter-conditioning, which involved repeated presentations of the same CS-US pairings. Yet, a central hypothesis of the propositional account is that repetition is not necessary for instruction effects on implicit evaluations (De Houwer, 2006). Consistent with this hypothesis, studies on instructed extinction have shown that verbal instructions can entirely eliminate conditioned fear responses (e.g., Mallan, Sax, & Lipp, 2009). In these studies, extinction instructions were provided only once and not repeated over multiple trials. Moreover, because dual-process accounts assume that repeated verbal information can have associative effects (see Gawronski & Bodenhausen, 2006; Rydell & McConnell, 2006), a repetition of counter-instructions would lead to same predictions (i.e., a reversal of EC effect on implicit evaluations), and thus blur the theoretical distinctions that are central for the current research. From this perspective, a comparison of single-shot verbal instructions and multiple trials of CS-US pairings represents the most diagnostic test of the two accounts.

To explain some unexpected findings in the EC literature, proponents of propositional accounts have argued that implicit evaluations might be less sensitive to new information than explicit evaluations (e.g., Zanon, De Houwer, Gast, & Smith, 2014). This assumption helps to reconcile propositional accounts with the asymmetric effect of counter-instructions on explicit and implicit evaluations. However, it fails to explain why EC effects on implicit evaluations were fully reversed as a result of counter-conditioning but not counter-instructions. Of course, propositional accounts might be able to explain the latter asymmetry by means of additional assumptions. Yet, even if a post hoc explanation can be generated for this outcome, propositional accounts seem inferior to dual-process accounts, which predict the full set of findings in an a priori fashion (cf. Gawronski & Bodenhausen, 2015).

Another potential argument against our conclusion is that the current studies used a repeated-measurement design in which CS evaluations were measured before and after the manipulation of counter-conditioning and counter-instructions. Because the evaluative priming task included pairings of the CSs with positive and negative target words, one could argue that these pairings were functionally similar to a counter-conditioning procedure. Although it is possible that the prime-target pairings in the evaluative priming task have such a "counter-conditioning" effect, it does not explain the obtained asymmetry between counter-conditioning and counter-instructions in reversing EC effects on explicit and implicit evaluations. If anything, the prime-target pairings in the evaluative priming task should have contributed to a reversal of EC effects in the counter-instructions condition. Moreover, because effects of prime-target pairings should have been equal in the two reversal conditions, they fail to explain the current finding that EC effects on implicit evaluations were reversed only in response to counter-conditioning but not counter-instructions. Again, although propositional accounts might be able to explain the latter outcome by means of additional assumptions, they seem inferior to dual-process accounts, which predict the full set of findings in an a priori fashion (cf. Gawronski & Bodenhausen, 2015).

Finally, it should be noted that the measures of implicit evaluation differ in terms of their underlying mechanisms, which can lead to different outcomes of the same experimental manipulation (e.g., Deutsch & Gawronski, 2009). To provide further evidence for our conclusion, future research should examine whether the current findings generalize to other measures of implicit evaluation.

Conclusion

The main goal of the current research was to test competing predictions about the effectiveness of counter-conditioning and counter-instructions in reversing EC effects on explicit and implicit evaluations. Although both procedures effectively reversed EC effects on explicit evaluations, only directly experienced CS-US pairings reversed EC effects on implicit evaluations. These findings question the functional equivalence of counter-conditioning and counter-instructions hypothesized by single-process propositional accounts. Yet, they support dual-process accounts, suggesting that associative and propositional processes jointly contribute to EC effects.

Acknowledgments

We thank Allison DeJong, Ashley Diaz, Stephanie Gonzalez, and Shama Momin for their help in collecting the data.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by grant 341601-2013 from the Natural Sciences and Engineering Research Council of Canada to the second author and by grant UMO-2015/18/E/HS6/00765 from Polish National Science Center to the third author.

Note

1. We report all measures, all conditions, and all data exclusions. All data were collected in one shot without intermittent statistical analyses. The predetermined sample size was set to 100 participants for both studies. Due to excessive sign-ups at the end of the semester, Experiment 2 includes a somewhat larger sample. Based on the meta-analytic effect size of $d = .52$ reported by Hofmann, De Houwer, Perugini, Baeyens, and Crombez (2010), a sample size of 100 provides a power of .999 to detect a significant EC effect. To obtain a power of .80 in detecting a significant difference between counter-conditioning versus counter-instructions in reversing EC effects, the effect size of their differential effectiveness would have to be $d = .57$ with a sample size of 100.

Supplemental Material

The supplemental material is available in the online version of the article.

References

De Houwer, J. (2006). Using the Implicit Association Test does not rule out an impact of conscious propositional knowledge on evaluative conditioning. *Learning and Motivation*, *37*, 176–187.

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 learning as an alternative for association formation models. *Learning and Behavior*, *37*, 1–20.

De Houwer, J. (2014). Why a propositional single-process model of associative learning deserves to be defended. In J. W. Sherman, B. Gawronski, & Y. Trope (Eds.), *Dual-process theories of the social mind* (pp. 530–541). New York, NY: Guilford Press.

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.

Deutsch, R., & Gawronski, B. (2009). When the method makes a difference: Antagonistic effects on "automatic evaluations" as a function of task characteristics of the measure. *Journal of Experimental Social Psychology*, *45*, 101–114.

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.

Gast, A., & De Houwer, J. (2012). Evaluative conditioning without directly experienced pairings of the conditioned and the unconditioned stimuli. *The Quarterly Journal of Experimental Psychology*, *65*, 1657–1674.

Gast, A., & De Houwer, J. (2013). The influence of extinction and counterconditioning instructions on evaluative conditioning effects. *Learning and Motivation*, *44*, 312–325.

Gast, A., Gawronski, B., & De Houwer, J. (2012). Evaluative conditioning: Recent developments and future directions. *Learning and Motivation*, *43*, 79–88.

Gawronski, B., Balas, R., & Creighton, L. A. (2014). Can the formation of conditioned attitudes be intentionally controlled? *Personality and Social Psychology Bulletin*, *40*, 419–432.

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., & Bodenhausen, G. V. (2011). The associative-propositional evaluation model: Theory, evidence, and open questions. *Advances in Experimental Social Psychology*, *44*, 59–127.

Gawronski, B., & Bodenhausen, G. V. (2015). Theory evaluation. In B. Gawronski & G. V. Bodenhausen (Eds.), *Theory and explanation in social psychology* (pp. 3–23). New York, NY: Guilford Press.

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.

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

Hu, X., Gawronski, B., & Balas, R. (2017). Propositional versus dual-process accounts of evaluative conditioning: I. The effects of co-occurrence and relational information on implicit and explicit evaluations. *Personality and Social Psychology Bulletin*, *43*, 17–32.

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

Jones, C. R., Olson, M. A., & Fazio, R. H. (2010). Evaluative conditioning: The "how" question. *Advances in Experimental Social Psychology*, *43*, 205–255.

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-7). Gainesville: University of Florida.

Mallan, K. M., Sax, J., & Lipp, O. V. (2009). Verbal instruction abolishes fear conditioned to racial out-group faces. *Journal of Experimental Social Psychology*, *45*, 1303–1307.

Mitchell, C. J., De Houwer, J., & Lovibond, P. F. (2009). The propositional nature of human associative learning. *Behavioral and Brain Sciences*, *32*, 183–198.

Rydell, R. J., & McConnell, A. R. (2006). Understanding implicit and explicit attitude change: A systems of reasoning analysis. *Journal of Personality and Social Psychology*, *91*, 995–1008.

Sweldens, S., Van Osselaer, S., & Janiszewski, C. (2010). Evaluative conditioning procedures and the resilience of conditioned brand attitudes. *Journal of Consumer Research*, *37*, 473–489.

Wentura, D., & Degner, J. (2010). A practical guide to sequential priming and related tasks. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (pp. 95–116). New York, NY: Guilford Press.

Zanon, R., De Houwer, J., Gast, A., & Smith, C. T. (2014). When does relational information influence evaluative conditioning? *The Quarterly Journal of Experimental Psychology*, *67*, 2105–2122.

Author Biographies

Xiaoqing Hu is an Assistant Professor in the Department of Psychology at the University of Hong Kong, China. He is interested in evaluative learning, memory and social cognition.

Bertram Gawronski is a Professor in the Department of Psychology at the University of Texas at Austin, USA. He is interested in social cognition, attitudes, decision-making, and moral psychology.

Robert Balas is an Assistant Professor of Psychology at the Polish Academy of Sciences, Poland. He is interested in evaluative conditioning, attitude formation and change.

Handling Editor: Kate Ratliff