Personality and Social Psychology Bulletin
37(4) 557–569

DOI: 10.1177/0146167211400423
http://pspb.sagepub.com
SAGE

Are We Puppets on a String? Comparing the Impact of Contingency and Validity on Implicit and Explicit Evaluations

Kurt R. Peters[1] **and Bertram Gawronski**[1]

Abstract

Research has demonstrated that implicit and explicit evaluations of the same object can diverge. Explanations of such dissociations frequently appeal to dual-process theories, such that implicit evaluations are assumed to reflect object-valence contingencies independent of their perceived validity, whereas explicit evaluations reflect the perceived validity of object-valence contingencies. Although there is evidence supporting these assumptions, it remains unclear if dissociations can arise in situations in which object-valence contingencies are judged to be true or false during the learning of these contingencies. Challenging dual-process accounts that propose a simultaneous operation of two parallel learning mechanisms, results from three experiments showed that the perceived validity of evaluative information about social targets qualified both explicit and implicit evaluations when validity information was available immediately after the encoding of the valence information; however, delaying the presentation of validity information reduced its qualifying impact for implicit, but not explicit, evaluations.

Keywords

implicit cognition, attitudes, automatic processes, control, dual-process models

Received June 21, 2010; revision accepted November 12, 2010

In the run-up to the 2008 U.S. presidential election, as the competition between Barack Obama and John McCain intensified, *The New Yorker* magazine published a cover illustration depicting Obama as a terrorist occupying the Oval Office. The resulting uproar from Obama's supporters reflected a suspicion that the illustration could lead voters to form negative associations with Obama even if they rejected the depicted link to terrorism (Banaji, 2008). Negative political campaigning has likewise been criticized for exploiting the ease with which evaluative associations can be manipulated (e.g., Carraro, Gawronski, & Castelli, 2010), and the veracity of these associations matters, as they have been shown to predict voting behavior in undecided voters independent of conscious beliefs (Galdi, Arcuri, & Gawronski, 2008; Payne et al., 2010). These examples suggest that learning might involve more than just the formation of beliefs, such that evaluative associations might be formed independent of, and even despite, conscious assessment of their validity. The unnerving implication is that we may be no more than "puppets on a string," helpless to resist being influenced by all the contingencies to which we are exposed in an information-saturated world.

The difference between associations and beliefs is supported by research in social cognition showing that it is possible for people to express divergent evaluations of the same object under different conditions (see Gawronski & Bodenhausen, 2006, for a review). For instance, explicit evaluations expressed under conditions of controlled processing often diverge from implicit evaluations expressed under conditions of automatic processing (Bargh, 1994). These results suggest that evaluative responses do not always reflect conscious beliefs about an object and that automatically activated associations may influence evaluative responses under suboptimal processing conditions.

In line with suspicions about the effects of media influence, evaluative dissociations are often explained by appeal to dual-process theories of learning, which posit two learning processes that may operate in parallel (e.g., Gawronski & Bodenhausen, 2006; Rydell & McConnell, 2006; Strack & Deutsch, 2004). On this account, explicit evaluations are the product of a *belief-based learning* process, which qualifies

[1]University of Western Ontario, London, ON, Canada

Corresponding Author:
Kurt R. Peters, University of Western Ontario, Department of Psychology, Social Science Centre, London, Ontario N6A 5C2, Canada
Email: kpeter3@uwo.ca

the evaluation implied by object-valence contingencies by the perceived validity of these contingencies. Implicit evaluations, in contrast, are thought to be the product of a *contingency-based learning* process, which encodes contingencies independent of their perceived validity and hence is sensitive to the dominant valence associated with an object. Because the two learning processes are assumed to operate in parallel, evaluative dissociations may already occur at the time of learning if the evaluation implied by observed contingencies is qualified by conscious beliefs about its validity.

Indeed, evidence for dual-process theories of learning seems compelling, such that evaluative dissociations may often arise at the time of learning about an object. For example, Rydell, McConnell, Mackie, and Strain (2006) employed a learning procedure in which participants had to guess the validity of valenced behavioral descriptions about a social target named Bob. Preceding the display of Bob's photograph on each trial, participants were subliminally presented with a prime word whose valence was opposite to the evaluation implied by the validity of the behavioral descriptions. Rydell et al. found that explicit evaluations of Bob reflected the valence of the valid behavioral descriptions whereas implicit evaluations of Bob reflected the valence of the subliminal primes.

Although findings such as these provide evidence that implicit and explicit evaluations are differentially sensitive to contingency-based versus belief-based learning processes, they remain silent about whether the two learning mechanisms can operate *simultaneously* on the basis of the *same* information (as is implied in the depiction of Obama as a terrorist). In particular, the currently available evidence is limited in answering this question given that demonstrations of learning-related dissociations typically involved multiple manipulations of an object's valence via distinct sources of information. For instance, in Rydell et al.'s (2006) study, contingency-based learning was driven by the subliminal primes that preceded the presentation of Bob, whereas belief-based learning was driven by the validity of the behavioral descriptions that followed the presentation of Bob. Moreover, the contingencies established through the priming manipulation may not be subject to conscious qualification if they are learned outside awareness. Consequently, it remains unclear if evaluative dissociations can arise in situations in which the evaluation implied by contingencies can itself be assessed as true or false. Such conditions are more consistent with the examples of media influence described above, where the observer is frequently confronted with information that is immediately perceived to be invalid.

The present experiments aim to address this question by directly manipulating the perceived validity of object-valence contingencies in a single learning episode. Based on dual-process theories that propose two parallel learning mechanisms (e.g., Gawronski & Bodenhausen, 2006; Rydell & McConnell, 2006; Strack & Deutsch, 2004), we expected an evaluative dissociation in this situation because of the simultaneous operation of belief-based and contingency-based learning processes. Specifically, given evidence that the negation of an association has been shown to qualify explicit, but not implicit, evaluations (Deutsch, Gawronski, & Strack, 2006; Gregg, Seibt, & Banaji, 2006), evaluative dissociations were expected to arise when the valence most frequently associated with an object is perceived to be false. In these cases, explicit evaluations should reflect the perceived validity of observed contingencies, whereas implicit evaluations should reflect observed contingencies independent of their perceived validity.

To anticipate our findings, this prediction was *not* confirmed in our studies. Contrary to the assumption that contingency-based and belief-based learning mechanisms may operate simultaneously on the basis of the same information, we found that the perceived validity of evaluative information about social targets qualified both explicit and implicit evaluations when validity information was available during the learning of the valence information. The expected dissociations occurred only when the presentation of validity information was delayed, which reduced its qualifying effect on implicit, but not explicit, evaluations. Taken together, these results support accounts that explain evaluative dissociations in terms of *expression-related*, as opposed to *learning-related*, processes, such that dissociations can be explained by the rejection of previously formed associations during the course of generating a controlled evaluative response (see Hofmann, Gschwendner, Nosek, & Schmitt, 2005, for a review). Consequently, these findings pose a challenge to the view that evaluative dissociations may be explained by the simultaneous operation of two independent learning mechanisms on the basis of the same information.

Experiment 1

In the first experiment, we sought to test the basic question of whether an evaluative dissociation can arise during a single learning episode, in which the perceived validity of object-valence contingencies is directly manipulated. The experiment involved a social learning task that required participants to form impressions of four novel targets by reading valenced behavioral descriptions about each of them. To manipulate the perceived validity of the resulting contingencies, each behavioral description was immediately marked as either true or false. The four targets in the learning task thus differed according to the valence with which they were most frequently associated (positive vs. negative) as well as the "true" valence of each target. The key empirical question is whether an evaluative dissociation will arise when the "true" valence diverges from the valence most frequently associated with the target. Based on the assumption that contingency-based and belief-based learning mechanisms may operate simultaneously on the basis of the same information, we originally expected that explicit evaluations would reflect the observed contingencies qualified by their perceived validity, whereas implicit evaluations would reflect these contingencies without any qualification.

Method

Participants and design. A total of 28 undergraduate students (20 women, 8 men) participated in a study on impression formation for course credit. The experiment employed a 2 (dominant valence: 75% positive vs. 75% negative) × 2 (validity of dominant valence: true vs. false) × 2 (evaluation type: explicit vs. implicit) factorial design with all three variables varying within participants. The order of the two evaluation measures was counterbalanced across participants.

Learning task. Upon entering the lab, participants were seated at individual computers and signed informed consent documents. Participants then began the computerized learning task, in which they were asked to form accurate impressions of four different people based on minimal information about each of them. The learning paradigm consisted of a guessing task in which participants were sequentially presented with photographs of four men along with valenced behavioral descriptions of each of them. Participants' task was to guess the accuracy of each description with feedback provided immediately following each guess (i.e., "RIGHT!" or "WRONG!"; see Rydell et al., 2006). Feedback following guesses was 100% consistent for all targets so that there was no misleading feedback about the true valence of any target. Each participant thus learned the true valence of each target according to his or her own pattern of guesses. In addition, instructions preceding the task informed participants that, when a behavioral description turned out to be false, they should infer that the opposite of the implied evaluation was true (i.e., a positive description that turns out to be false implies a negative evaluation, and vice versa).

A total of 20 behavioral descriptions (adapted from Rydell et al., 2006) were presented for each of the four targets, including both positively valenced descriptions (e.g., "Mike lent money to a friend in financial trouble") and negatively valenced descriptions (e.g., "Mike cheated during a poker game"). The valence of the behavioral descriptions for each target was held in 3:1 proportion such that two targets were paired with 15 positive and 5 negative descriptions and two targets were paired with 5 positive and 15 negative descriptions. The "true" valence was varied orthogonally to the dominant valence of the behavioral descriptions so that either the positive descriptions were true and the negative descriptions were false, or vice versa. These two manipulations created four different impression-formation targets: (a) a target with mostly positive descriptions that were described as accurate, (b) a target with mostly negative descriptions that were described as accurate, (c) a target with mostly positive descriptions that were described as inaccurate, and (d) a target with mostly negative descriptions that were described as inaccurate. The particular mappings of the four photographs with the four experimental conditions were counterbalanced across participants.

With 20 behavioral descriptions presented for each of the four targets, the learning procedure consisted of a total of 80 trials presented to each participant in computer-randomized order. Each learning trial started with the presentation of a shoulder-up photograph of one of the four impression-formation targets, all of whom were young, White men, centered on the computer screen. At the same time, a valenced behavioral description was displayed below the picture. Participants were instructed to use two response keys on the keyboard to indicate their true–false guess on each trial. Upon making a response, the display was cleared and participants were given feedback about the validity of their guess. The feedback remained centered on the screen for 1,000 ms, followed by a 1,000-ms intertrial interval.

Measurement of explicit evaluations. Following the learning procedure, participants completed measures of explicit and implicit evaluations in counterbalanced order. To assess explicit evaluations, participants completed three self-report items (likeability, friendliness, and trustworthiness) for each of the four impression-formation targets in computer-randomized order. Responses for all items were made on 7-point rating scales ranging from 1 (*not at all*) to 7 (*very much*).

Measurement of implicit evaluations. The affect misattribution procedure (AMP; Payne, Cheng, Govorun, & Stewart, 2005) was used to measure implicit evaluations of each of the four impression-formation targets. Each trial of the AMP was displayed in the following sequence: A fixation cross was presented for 1,000 ms, a valenced stimulus (i.e., a photograph of one of the four targets) for 75 ms, a blank screen for 125 ms, and a Chinese pictograph for 100 ms, and finally a pattern mask of black-and-white noise was presented. Participants were instructed that, upon presentation of the mask, they were to indicate how "visually pleasant" they found the preceding Chinese pictograph using two response keys on the keyboard signifying *less pleasant* and *more pleasant*. Following Payne et al. (2005), participants were told that the pictures appearing before the Chinese pictographs may bias responses and that they should try not to let the pictures influence their judgments. In all, 25 AMP trials were presented for each impression-formation target, resulting in a total of 100 trials presented in computer-randomized order. Participants were debriefed following completion of the dependent measures.

Results

Data preparation. The three self-report items were averaged to create an index of the explicit evaluation of each of the four targets (all Cronbach's αs > .56). To create an index of the implicit evaluation of each target, the proportion of *more pleasant* responses on the relevant AMP trials was calculated, which varied between 0% (*negative*) and 100% (*positive*).

Explicit and implicit evaluations. The primary analyses collapsed across the order of the two evaluation measures. To test the effects of validity information on explicit and implicit evaluations, indices of both explicit and implicit evaluations were standardized to obtain a common metric and then submitted to a 2 (dominant valence: positive vs. negative) ×

2 (validity of dominant valence: true vs. false) × 2 (evaluation type: explicit vs. implicit) repeated measures analysis of variance (ANOVA). Significant main effects were observed for valence, $F(1, 27) = 19.07$, $p < .001$, $\eta^2_p = .41$, and validity, $F(1, 27) = 11.77$, $p = .002$, $\eta^2_p = .30$. In addition, significant two-way interactions were observed between valence and validity, $F(1, 27) = 73.07$, $p < .001$, $\eta^2_p = .73$, and validity and evaluation type, $F(1, 27) = 25.89$, $p < .001$, $\eta^2_p = .49$. Finally, the three-way interaction among valence, validity, and evaluation type was significant, $F(1, 27) = 17.63$, $p < .001$, $\eta^2_p = .40$. Further inspection of this interaction suggests that the qualification of the Validity × Valence effect by evaluation type does not reflect the expected dissociation between explicit and implicit evaluations as a function of validity. Instead, the interaction simply reflects a slightly weaker effect size of the Valence × Validity crossover interaction for implicit evaluations, as described below.

To specify the obtained three-way interaction, the effects of the valence and validity manipulations were assessed separately for both explicit and implicit evaluations using raw scores for all analyses. With respect to explicit evaluations, significant main effects of valence, $F(1, 27) = 22.53$, $p < .001$, $\eta^2_p = .46$, and validity, $F(1, 27) = 64.57$, $p < .001$, $\eta^2_p = .63$, were observed, qualified by a significant two-way interaction, $F(1, 27) = 72.30$, $p < .001$, $\eta^2_p = .73$. As shown in Figure 1A, validity information influenced explicit evaluations as expected, such that explicit evaluations reflected the dominant valence when it was true and the opposite of the dominant valence when it was false. Paired samples *t* tests revealed that when the dominant valence was true, explicit evaluations favored positively described targets over negatively described targets, $t(27) = 9.08$, $p < .001$; but when the dominant valence was false, explicit evaluations favored negatively described targets over positively described targets, $t(27) = 6.04$, $p < .001$. Moreover, when the dominant valence of behavioral descriptions was positive, explicit evaluations were more positive when the validity feedback for the dominant information was true rather than false, $t(27) = 12.80$, $p < .001$; but when the dominant valence of behavioral descriptions was negative, explicit evaluations were more positive when the validity feedback for the dominant information was false rather than true, $t(27) = 4.28$, $p < .001$.

With respect to implicit evaluations, no main effects were significant, but the two-way interaction between valence and validity was significant, $F(1, 27) = 16.85$, $p < .001$, $\eta^2_p = .38$. As shown in Figure 1B, the interaction effect for implicit evaluations was identical to that obtained for explicit evaluations; that is, implicit evaluations reflected the dominant valence when it was true and the opposite of the dominant valence when it was false. Paired-samples *t* tests revealed that when the dominant valence was true, implicit evaluations favored positively described targets over negatively described targets, $t(27) = 4.03$, $p < .001$; but when the dominant valence was false, implicit evaluations favored negatively described targets over positively described targets, $t(27) = 2.66$, $p = .013$.

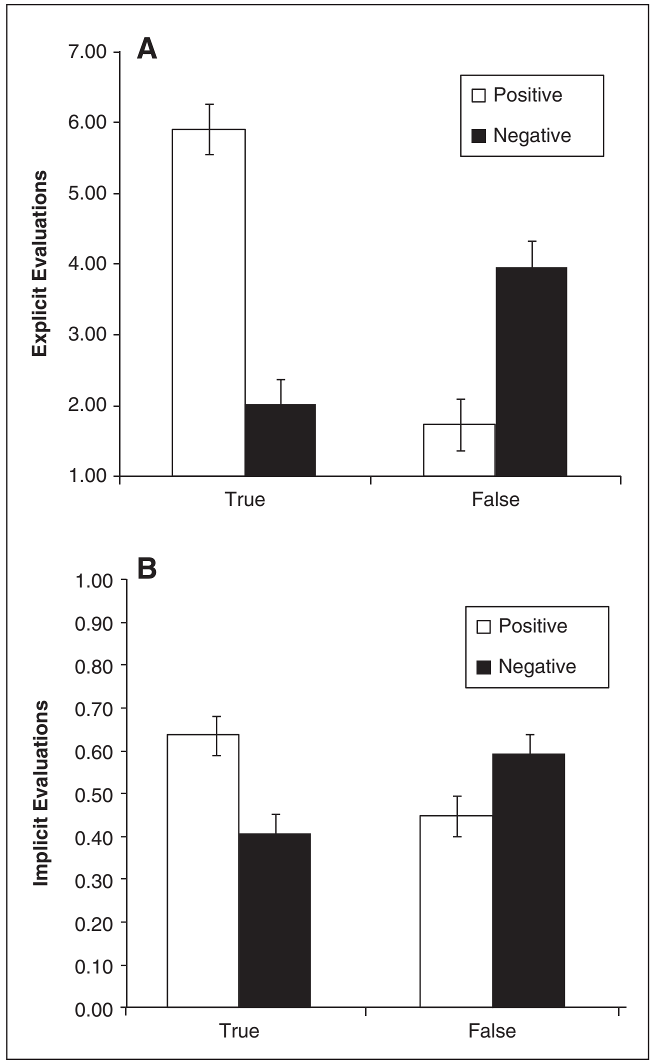

Figure 1. Top panel (Figure 1A): Explicit evaluations as a function of dominant valence (positive vs. negative) and validity of dominant valence (true vs. false), Experiment 1; bottom panel (Figure 1B): Implicit evaluations as a function of dominant valence (positive vs. negative) and validity of dominant valence (true vs. false), Experiment 1
Note: Error bars represent standard errors.

Moreover, when the dominant valence of behavioral descriptions was positive, implicit evaluations were more positive when the validity feedback for the dominant information was true rather than false, $t(27) = 3.16$, $p = .004$; but when the dominant valence of behavioral descriptions was negative, implicit evaluations were more positive when the validity feedback for the dominant information was false rather than true, $t(27) = 3.44$, $p = .002$.[1]

Discussion

Counter to our predictions, Experiment 1 revealed that perceived validity qualified the effect of object-valence contingencies for both explicit and implicit evaluations. When the dominant valence was true, explicit and implicit evaluations reflected the dominant valence, but when the dominant valence was false, explicit and implicit evaluations reflected the opposite of the dominant valence. Thus, no evaluative dissociation was observed when the perceived validity of observed contingencies was manipulated in a single learning episode. This pattern of results challenges the assumption that belief-based and contingency-based learning may operate simultaneously on the basis of the same information. Drawing on dual-process theories that propose the simultaneous operation of two parallel learning mechanisms, we originally expected that explicit evaluations would reflect the perceived validity of contingencies, whereas implicit evaluations would reflect contingencies independent of their perceived validity. This prediction was clearly disconfirmed in the current study. There is, however, a methodological concern with drawing this conclusion directly from the present results, which was addressed in the next experiment.

Experiment 2

The present experiment sought to rule out the concern that the absence of a dissociation in Experiment 1 resulted from inadequate measurement procedures. Recent evidence suggests that validity information pertaining to the primes can influence responses on the AMP (Deutsch, Kordts-Freudinger, Gawronski, & Strack, 2009), whereas Fazio's evaluative priming task (EPT; Fazio, Jackson, Dunton, & Williams, 1995) remains unaffected by validity information (Deutsch et al., 2006, 2009). Moreover, the AMP and the EPT have been shown to produce divergent effects of the same experimental manipulation under some conditions, suggesting that task-specific mechanisms may shape responses on these measures in a nontrivial manner (e.g., Deutsch & Gawronski, 2009; Gawronski, Cunningham, LeBel, & Deutsch, 2010). It would therefore be valuable to replicate the findings from Experiment 1 using an EPT to ensure that they are not unique to one measure of implicit evaluations but reflect a genuine effect that replicates across different measures of the same construct.

Method

Participants and design. A total of 45 undergraduate students (37 women, 8 men) participated in a study on impression formation for course credit. One participant was excluded from the analysis because of chance responding on the EPT (error rate > 40%). As with Experiment 1, this experiment employed a 2 (dominant valence: 75% positive vs. 75% negative) × 2 (validity of dominant valence: true vs. false) × 2 (evaluation type: explicit vs. implicit) factorial design with all three variables varying within participants. The order of the two evaluation measures was counterbalanced across participants.

Procedure. The learning task and the measure of explicit evaluations were identical to Experiment 1. Fazio et al.'s (1995) EPT was used to measure implicit evaluations of the four impression-formation targets. Each trial of the EPT was displayed in the following sequence: A fixation cross was presented for 500 ms, a valenced prime (i.e., a photograph of one of the four targets) was presented for 200 ms, and then a positive or negative target word (e.g., *paradise* or *poison*) was presented and remained on-screen until the participant indicated whether the word was positive or negative using one of two response keys on the keyboard. If the response was incorrect, "*ERROR!*" was displayed for 1,500 ms. An interval of 1,000 ms preceded the start of the next trial. According to Fazio et al. (1995), the affect elicited by the prime should facilitate evaluative decisions for valence-congruent target words but inhibit evaluative decisions for valence-incongruent target words, so that response latencies to the target words can be used to infer implicit evaluations of each impression-formation target. Each of the four targets served as a prime on 20 trials, split between 10 trials with negative and 10 trials with positive target words, creating a total of 80 trials presented in computer-randomized order. Participants were debriefed following completion of the dependent measures.

Results

Data preparation. Indices of the explicit evaluation of each of the four impression-formation targets were calculated as described in Experiment 1 (all Cronbach's αs > .83). In creating indices of the implicit evaluation of each of the four impression-formation targets, EPT trials with incorrect responses (5.1%) were excluded. To control for anticipations and outliers (Ratcliff, 1993), response cutoffs were employed to exclude trials with reaction times shorter than 300 ms or longer than 1,000 ms (8.7% of valid trials). Then, for each of the four primes in the EPT, the mean reaction time to trials with positive target words was subtracted from that for trials with negative target words, so that higher scores reflect a relatively more positive implicit evaluation of the target (see Wentura & Degner, 2010).

Explicit and implicit evaluations. The order of the two evaluation measures had no effect, so analyses collapsed across this factor. To test the effects of validity information on explicit and implicit evaluations, indices of both explicit and implicit evaluations were standardized and submitted to a 2 (dominant valence: 75% positive vs. 75% negative) × 2 (validity of dominant valence: true vs. false) × 2 (evaluation type: explicit vs. implicit) repeated measures ANOVA. Significant two-way interactions were observed between valence and validity, $F(1, 43) = 358.08$, $p < .001$, $\eta^2_p = .89$, and valence and evaluation type, $F(1, 43) = 4.53$, $p = .039$, $\eta^2_p = .10$. The three-way interaction among valence, validity, and evaluation type was also significant, $F(1, 43) = 270.53$, $p < .001$, $\eta^2_p = .86$. No other effects were significant. As with Experiment 1,

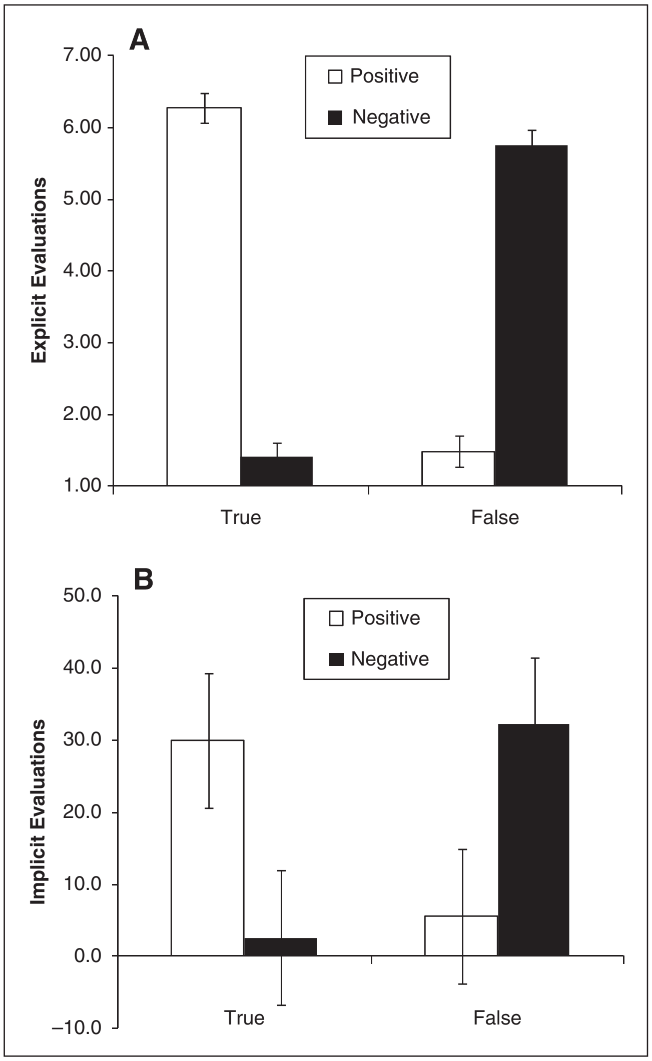

Figure 2. Top panel (Figure 2A): Explicit evaluations as a function of dominant valence (positive vs. negative) and validity of dominant valence (true vs. false), Experiment 2; bottom panel (Figure 2B): Implicit evaluations as a function of dominant valence (positive vs. negative) and validity of dominant valence (true vs. false), Experiment 2
Note: Error bars represent standard errors.

the qualification of the Validity × Valence effect by evaluation type does not reflect the expected dissociation between explicit and implicit evaluations. Instead, the interaction reflects a slightly weaker effect size of the Valence × Validity crossover interaction for implicit evaluations, as described below.

To specify the obtained three-way interaction, the effects of the valence and validity manipulations were assessed separately for both explicit and implicit evaluations, using raw scores for all analyses. With respect to explicit evaluations, significant main effects were observed for valence, $F(1, 43) = 8.10$, $p = .007$, $\eta^2_p = .16$, and for validity, $F(1, 43) = 5.39$, $p = .025$, $\eta^2_p = .11$, qualified by a significant two-way interaction, $F(1, 43) = 445.49$, $p < .001$, $\eta^2_p = .91$. As shown in Figure 2A, validity information influenced explicit evaluations as expected, such that explicit evaluations reflected the dominant valence when it was true and the opposite of the dominant valence when it was false. Paired samples *t* tests revealed that when the dominant valence was true, explicit evaluations favored positively described targets over negatively described targets, $t(43) = 21.60$, $p < .001$; but when the dominant valence was false, explicit evaluations favored negatively described targets over positively described targets, $t(43) = 16.67$, $p < .001$. Moreover, when the dominant valence of behavioral descriptions was positive, explicit evaluations were more positive when the validity feedback for the dominant information was true rather than false, $t(43) = 21.98$, $p < .001$; but when the dominant valence of behavioral descriptions was negative, explicit evaluations were more positive when the validity feedback for the dominant information was false rather than true, $t(43) = 17.17$, $p < .001$.

With respect to implicit evaluations, no main effects were significant, but the two-way interaction between valence and validity was significant, $F(1, 43) = 8.31$, $p = .006$, $\eta^2_p = .16$. As shown in Figure 2B, the effect of validity information on implicit evaluations was identical to its effect on explicit evaluations; that is, implicit evaluations reflected the dominant valence when it was true and the opposite of the dominant valence when it was false. Paired samples *t* tests revealed that when the dominant valence was true, implicit evaluations favored positively described targets over negatively described targets, $t(43) = 2.15$, $p = .037$; but when the dominant valence was false, implicit evaluations favored negatively described targets over positively described targets, $t(43) = 2.29$, $p = .027$. Moreover, when the dominant valence of behavioral descriptions was positive, implicit evaluations were more positive when the validity feedback for the dominant information was true rather than false, $t(43) = 2.11$, $p = .041$; but when the dominant valence of behavioral descriptions was negative, implicit evaluations were more positive when the validity feedback for the dominant information was false rather than true, $t(43) = 2.42$, $p = .020$.

Discussion

The findings of Experiment 2 replicated those of Experiment 1 using a measure of implicit evaluations less sensitive to validity information (Deutsch et al., 2006, 2009). Once again, perceived validity qualified the effect of object-valence contingencies for both explicit and implicit evaluations. These results rule out the concern that the absence of dissociation in Experiment 1 was the result of suboptimal measurement procedures.

Experiment 3

Experiments 1 and 2 suggest that during a single learning episode, the perceived validity of object-valence contingencies influences both explicit and implicit evaluations. In other words, it appears that in situations that involve exposure to information that is considered invalid, it is possible to exercise control over what is learned. There is, however, evidence for evaluative dissociations arising from asymmetric influences of validity information on explicit and implicit evaluations. For example, Gregg et al. (2006) found that both explicit and implicit evaluations initially reflected the valence of behavioral descriptions of two novel groups but that subsequently acquired information about the validity of these descriptions qualified explicit, but not implicit, evaluations. This finding is at odds with the results of the preceding experiments, in which validity information qualified both explicit and implicit evaluations.

An important factor that may account for the difference between the two findings is the time at which validity information was provided. Whereas in our studies validity information was available during the learning of the behavioral descriptions, Gregg et al.'s (2006) study included a substantial delay between the initial learning of the behavioral descriptions and the subsequent presentation of validity information. Thus, counter to the notion of *learning-related* dissociations because of the simultaneous operation of two distinct learning mechanisms on the basis of the same information, Gregg et al.'s (2006) findings are better described as a case of *expression-related* dissociations. Such dissociations occur when information that has been stored in memory at an earlier time is later learned to be invalid. In such cases, newly acquired validity information may be unable to erase previously formed associations from memory, even though these associations are rejected as invalid in the course of expressing an explicit evaluative judgment (Fazio, 2007; Gawronski & Bodenhausen, 2006). As a result, newly acquired validity information will influence explicit, but not implicit, evaluations. In fact, many examples of evaluative dissociations can be parsimoniously explained by the subsequent rejection of previously learned information without assuming a simultaneous operation of two independent learning mechanisms (see Hofmann et al., 2005, for a review).

If this interpretation is correct, then it should be possible to create an evaluative dissociation using the present experimental paradigm by manipulating the delay between the presentation of the behavioral descriptions and information about their validity. Experiments 1 and 2 showed that, when validity information is available during the learning of the behavioral descriptions, perceived validity produces equivalent effects on explicit and implicit evaluations. On the other hand, if the presentation of validity information is delayed, it must be applied to existing associations post hoc, presumably qualifying explicit, but not implicit, evaluations. This pattern would be consistent with the idea that evaluative dissociations arising from perceived validity are the result of expression-related, rather than learning-related, processes. Experiment 3 was designed to test this hypothesis.

Method

Participants and design. A total of 218 undergraduate students (159 women, 59 men; mean age = 22.03) participated in a study on impression formation for course credit. Data from 14 participants were unusable because of a programming error, and another 15 participants were excluded because of chance responding on the EPT (error rates > 40%). The final sample consisted of 189 students (139 women, 50 men). The experiment employed a 2 (valence: positive vs. negative) × 2 (validity of valence: true vs. false) × 2 (evaluation type: explicit vs. implicit) × 2 (validity timing: short delay vs. long delay) factorial design with the first three variables varying within participants and the last varying between participants. The order of the two evaluation measures was counterbalanced across participants.

Learning procedure. The learning procedure employed in Experiment 3 was broadly similar to that used in Experiments 1 and 2, with a few important differences. First, a more detailed cover story was provided, which framed the learning procedure in terms of learning about coworkers at a new job based on secondhand comments (adapted from Gawronski & Walther, 2008). To strengthen the overall effect of valence during the learning procedure, the four impression-formation targets were paired with 100% positive or 100% negative behavioral descriptions. The guessing component of the procedure was therefore dropped, and the learning task was instead introduced as a slideshow that required only that participants attend to the information presented. Furthermore, because 100% consistent behavioral descriptions should be easily learned, only 5 learning trials were displayed for each target, for a total of 20 learning trials presented in computer-randomized order. New positive and negative behavioral descriptions were created to conform with the "workplace" cover story (adapted from Gawronski, Walther, & Blank, 2005).

To test the effects of immediate versus delayed presentation of validity information on explicit and implicit evaluations, the delivery of validity information during the learning procedure was manipulated to be either (a) interleaved with the learning trials in the short-delay condition or (b) presented after all learning trials had finished in the long-delay condition. Instructions prior to the learning task in the short-delay condition informed participants that some behavioral descriptions would turn out to be false and that in these cases they should infer that the opposite of the implied evaluation was true. In the long-delay condition, instructions prior to the learning task informed participants that some behavioral descriptions would turn out to be false but that they should initially assume that all of the descriptions are true.

Each learning trial in the short-delay condition (similar to Experiments 1 and 2) began with the presentation of a photograph of one of the four targets together with a valenced behavioral description. After 3,000 ms, validity information was presented just below the behavioral description and remained on-screen for another 3,000 ms. A 1,500-ms intertrial interval preceded the start of the next learning trial. In the long-delay condition, each learning trial began with the presentation of a photograph of one of the four targets together with a valenced behavioral description. This information remained on-screen for 6,000 ms, and a 1,500-ms intertrial interval preceded the start of the next trial. The total duration of the 20-trial slideshow in both conditions was 150 s.

Following completion of the slideshow in the short-delay condition, participants were asked to take a moment to integrate the behavioral descriptions with the validity information to arrive at a clear impression of each target and to proceed to the next component of the study at their own pace. In the long-delay condition, participants were told that the behavioral descriptions for two of the targets were all true whereas the behavioral descriptions for the other two targets were all false. A photograph of each target and the validity of the descriptions associated with that target (i.e., "TRUE COMMENTS" or "FALSE COMMENTS") were displayed on one screen to make this clear. Participants were asked to take their time to arrive at a clear impression of each target in light of the new validity information. In both conditions, the valence and validity of the four targets were crossed to produce a positive–true, positive–false, negative–true, and negative–false target. The particular mappings of the four photographs with the four experimental conditions were counterbalanced across participants.

Measurement of explicit and implicit evaluations. The measures of explicit evaluations of each target were identical to those used in Experiments 1 and 2. Implicit evaluations of each target were assessed using an EPT identical to that used in Experiment 2, except that the total number of trials was doubled to 160. Participants were debriefed following completion of the dependent measures.

Results

Data preparation. Indices of the explicit evaluation of each of the four impression-formation targets were calculated as described in Experiment 1 (all Cronbach's αs > .90). In creating indices of the implicit evaluation of each of the four impression-formation targets, EPT trials with incorrect responses (4.3%) were excluded. Response cutoffs were also employed to exclude trials with reaction times shorter than 300 ms or longer than 1,000 ms (7.3% of valid trials). Calculation of the implicit indices from the EPT scores followed the procedure described in Experiment 2.

Explicit and implicit evaluations. The order of the two evaluation measures had no effect, so analyses collapsed across this factor. To test the effects of delayed validity information on explicit and implicit evaluations, indices of both explicit and implicit evaluations were standardized and submitted to a 2 (valence: positive vs. negative) × 2 (validity of valence: true vs. false) × 2 (evaluation type: explicit vs. implicit) × 2 (validity timing: short delay vs. long delay) mixed-model ANOVA with repeated measures on the first three factors. Significant main effects were observed for valence, $F(1, 187) = 54.48$, $p < .001$, $\eta^2_p = .23$, and validity, $F(1, 187) = 13.28$, $p < .001$, $\eta^2_p = .07$. In addition, significant two-way interactions were observed between valence and validity, $F(1, 187) = 618.97$, $p < .001$, $\eta^2_p = .77$; between valence and timing, $F(1, 187) = 11.61$, $p = .001$, $\eta^2_p = .06$; between valence and evaluation type, $F(1, 187) = 11.58$, $p = .001$, $\eta^2_p = .06$; and between validity and evaluation type, $F(1, 187) = 4.64$, $p = .033$, $\eta^2_p = .02$. Significant three-way interactions were observed among valence, validity, and timing, $F(1, 187) = 18.91$, $p < .001$, $\eta^2_p = .09$; among valence, validity, and evaluation type, $F(1, 187) = 532.18$, $p < .001$, $\eta^2_p = .74$; and among validity, timing, and evaluation type, $F(1, 187) = 7.27$, $p = .008$, $\eta^2_p = .04$. Finally, and most relevant to the current hypothesis, a significant four-way interaction was observed, $F(1, 187) = 14.07$, $p < .001$, $\eta^2_p = .07$, indicating that the effects of valence and validity on explicit and implicit evaluations were differentially moderated by the timing of validity information. To specify the particular nature of this interaction, analyses of explicit and implicit evaluations are reported separately for each of the two validity timing conditions.

Evaluations under short-delay validity timing. The condition involving a short delay before the presentation of validity information is conceptually identical to the design employed in Experiments 1 and 2, and analyses proceed similarly. To test the effects of valence and validity feedback on explicit and implicit evaluations in the short-delay condition, standardized indices of explicit and implicit evaluations were submitted to a 2 (valence: positive vs. negative) × 2 (validity of valence: true vs. false) × 2 (evaluation type: explicit vs. implicit) repeated measures ANOVA. A significant main effect was observed for valence, $F(1, 102) = 10.51$, $p = .002$, $\eta^2_p = .09$. In addition, significant two-way interactions were observed between valence and validity, $F(1, 102) = 561.63$, $p < .001$, $\eta^2_p = .85$, and between valence and evaluation type, $F(1, 102) = 4.14$, $p = .044$, $\eta^2_p = .04$. Finally, the three-way interaction among valence, validity, and evaluation type was significant, $F(1, 102) = 393.56$, $p < .001$, $\eta^2_p = .79$. No other effects were significant. As with Experiments 1 and 2, the qualification of the Validity × Valence interaction by evaluation type does not reflect the expected dissociation between explicit and implicit evaluations. Instead, the interaction reflected a slightly weaker effect size of the Valence × Validity crossover interaction for implicit evaluations, as described below.

The effects of the valence and validity manipulations were assessed separately for both explicit and implicit evaluations using raw scores for all analyses. With respect to explicit evaluations, a significant main effect of valence was observed, $F(1, 102) = 17.98$, $p < .001$, $\eta^2_p = .15$, qualified by a significant two-way interaction between valence and validity, $F(1, 102) = 763.14$, $p < .001$, $\eta^2_p = .88$. As shown in Figure 3A, validity information influenced explicit evaluations as expected, such that explicit evaluations reflected the valence

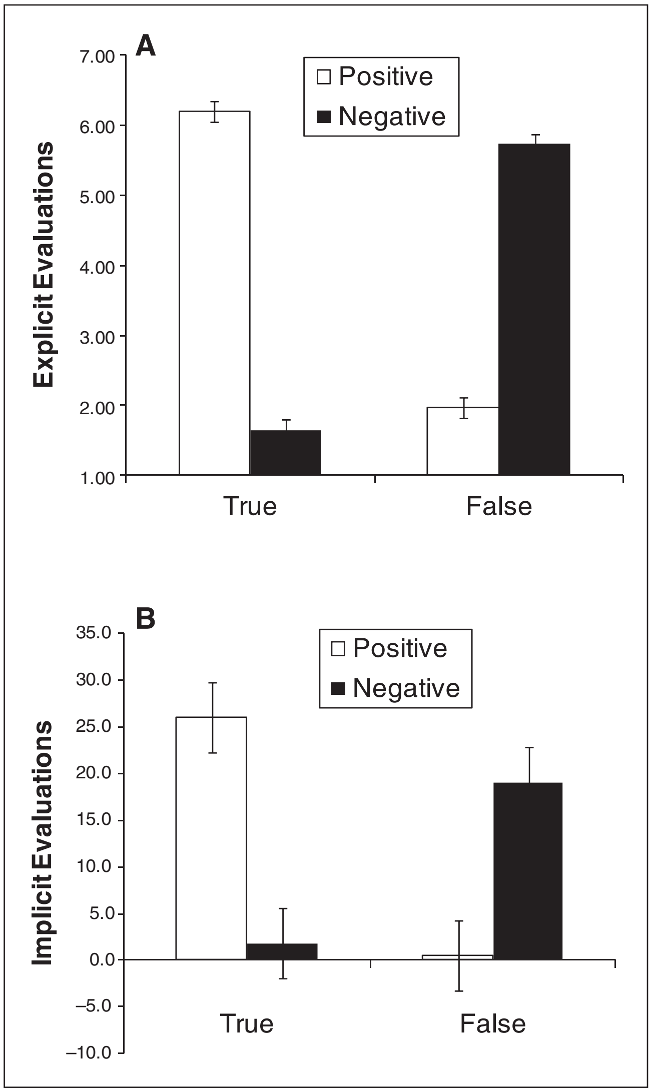

Figure 3. Top panel (Figure 3A): Explicit evaluations as a function of dominant valence (positive vs. negative) and validity of dominant valence (true vs. false) with short-delay validity feedback, Experiment 3; bottom panel (Figure 3B): Implicit evaluations as a function of dominant valence (positive vs. negative) and validity of dominant valence (true vs. false) with short-delay validity feedback, Experiment 3
Note: Error bars represent standard errors.

of behavioral descriptions when they were true and the opposite valence when they were false. Paired samples *t* tests revealed that when the behavioral descriptions were true, explicit evaluations favored positively described targets over negatively described targets, $t(102) = 26.27$, $p < .001$; but when the behavioral descriptions were false, explicit evaluations favored negatively described targets over positively described targets, $t(102) = 20.92$, $p < .001$. Moreover, when the valence of behavioral descriptions was positive, explicit evaluations were more positive when the validity information was true rather than false, $t(102) = 21.52$, $p < .001$; but when the valence of behavioral descriptions was negative, explicit evaluations were more positive when the validity information was false rather than true, $t(102) = 25.08$, $p < .001$.

With respect to implicit evaluations, no main effects were significant, but the two-way interaction between Valence and Validity was significant, $F(1, 102) = 32.37$, $p < .001$, $\eta^2_p = .24$. As shown in Figure 3B, the pattern of the Valence × Validity interaction was identical to that obtained for explicit evaluations. Specifically, under quick validity feedback, implicit evaluations reflected the dominant valence when it was true and the opposite valence when it was false. Paired samples *t* tests revealed that when the behavioral descriptions were true, implicit evaluations favored positively described targets over negatively described targets, $t(102) = 4.62$, $p < .001$; but when the behavioral descriptions were false, implicit evaluations favored negatively described targets over positively described targets, $t(102) = 3.44$, $p = .001$. Moreover, when the valence of behavioral descriptions was positive, implicit evaluations were more positive when the validity information was true rather than false, $t(102) = 5.11$, $p < .001$; but when the valence of behavioral descriptions was negative, implicit evaluations were more positive when the validity information was false rather than true, $t(102) = 3.72$, $p < .001$.

Taken together, these results replicate the findings of Experiments 1 and 2: When the delay between the presentation of valence and validity information was short, both explicit and implicit evaluations reflected the qualification of the behavioral descriptions by their perceived validity.

Evaluations under long-delay validity timing. To test for a potential dissociation between explicit and implicit evaluations in the long-delay condition, standardized indices of explicit and implicit evaluations were submitted to a 2 (valence: positive vs. negative) × 2 (validity of valence: true vs. false) × 2 (evaluation type: explicit vs. implicit) repeated measures ANOVA. Significant main effects were observed for valence, $F(1, 85) = 44.15$, $p < .001$, $\eta^2_p = .34$, and validity, $F(1, 85) = 11.46$, $p = .001$, $\eta^2_p = .12$. In addition, significant two-way interactions were observed between valence and validity, $F(1, 85) = 161.51$, $p < .001$, $\eta^2_p = .66$; between valence and evaluation type, $F(1, 85) = 6.99$, $p = .010$, $\eta^2_p = .08$; and between validity and evaluation type, $F(1, 85) = 9.02$, $p = .004$, $\eta^2_p = .10$. Finally, the three-way interaction among valence, validity, and evaluation type was significant, $F(1, 85) = 172.05$, $p < .001$, $\eta^2_p = .67$. No other effects were significant. In this case, contrary to the results under quick validity timing, the qualification of the Validity × Valence effect by evaluation type does reflect a dissociation between explicit and implicit evaluations, as described below.

The effects of the valence and validity manipulations were assessed separately for both explicit and implicit evaluations using raw scores for all analyses. With respect to explicit evaluations, significant main effects were observed for valence, $F(1, 85) = 33.38$, $p < .001$, $\eta^2_p = .28$, and validity,

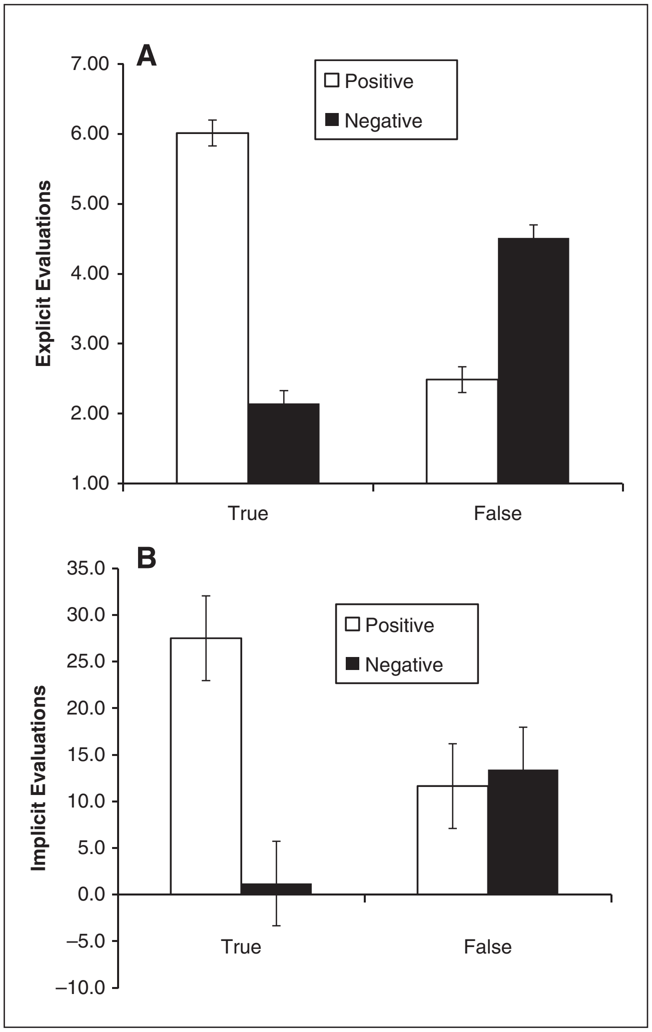

Figure 4. Top panel (Figure 4A): Explicit evaluations as a function of dominant valence (positive vs. negative) and validity of dominant valence (true vs. false) with long-delay validity feedback, Experiment 3; bottom panel (Figure 4B): Implicit evaluations as a function of dominant valence (positive vs. negative) and validity of dominant valence (true vs. false) with long-delay validity feedback, Experiment 3
Note: Error bars represent standard errors.

$F(1, 85) = 19.60$, $p < .001$, $\eta^2_p = .19$, qualified by a significant two-way interaction, $F(1, 85) = 254.95$, $p < .001$, $\eta^2_p = .75$. As shown in Figure 4A, validity information influenced explicit evaluations as expected, such that explicit evaluations reflected the valence of the behavioral descriptions when they were true and the opposite valence when they were false. Paired samples *t* tests revealed that when the behavioral descriptions were true, explicit evaluations favored positively described targets over negatively described targets, $t(85) = 16.72$, $p < .001$; but when the behavioral descriptions were false, explicit evaluations favored negatively described targets over positively described targets, $t(85) = 7.93$, $p < .001$. Moreover, when the valence of behavioral descriptions was positive, explicit evaluations were more positive when the validity information was true rather than false, $t(85) = 16.18$, $p < .001$; but when the valence of behavioral descriptions was negative, explicit evaluations were more positive when the validity information was false rather than true, $t(85) = 10.11$, $p < .001$.

With respect to implicit evaluations, a significant main effect of valence was observed, $F(1, 85) = 9.75$, $p = .002$, $\eta^2_p = .10$, qualified by a significant two-way interaction between valence and validity, $F(1, 85) = 9.55$, $p = .003$, $\eta^2_p = .10$. No other effects were significant. As shown in Figure 4B, the main effect of valence revealed that implicit evaluations of the positive targets were on average more positive than evaluations of the negative targets. This main effect of valence was qualified, however, by the validity of the behavioral descriptions, such that implicit evaluations reflected the valence of the behavioral descriptions when they turned out to be true, but this effect was only attenuated (rather than reversed) when the behavioral descriptions turned out to be false. Paired samples *t* tests revealed that when the behavioral descriptions were true, implicit evaluations favored positively described targets over negatively described targets, $t(85) = 38.88$, $p < .001$; but when the behavioral descriptions were false, implicit evaluations of negatively described targets were not significantly different from implicit evaluations of positively described targets, $t(85) = 0.31$, $p = .756$. Moreover, when the valence of behavioral descriptions was positive, implicit evaluations were more positive when the validity information was true rather than false, $t(85) = 2.37$, $p = .020$; but when the valence of behavioral descriptions was negative, implicit evaluations were more positive when the validity information was false rather than true, $t(85) = 2.25$, $p = .027$.

Thus, when the delay between the presentation of valence and validity information was relatively long, explicit and implicit evaluations became dissociated such that explicit evaluations reflected the full qualification of the behavioral descriptions by the validity information, but effects on implicit evaluations were merely attenuated.

Discussion

The results of Experiment 3 support the hypothesis that evaluative dissociations may arise when the acquisition of validity information is delayed. When validity information was available during the learning of evaluative information, it qualified both explicit and implicit evaluations, replicating the results of Experiments 1 and 2. When the presentation of validity information was delayed, however, its impact was significantly reduced for implicit, but not explicit, evaluations. The current results thus imply a boundary condition on the emergence of evaluative dissociations, such that validity information may qualify both explicit and implicit evaluations when it is available during the acquisition of evaluative information; with the passage of time, however, changes in the perceived validity of previously acquired information may still qualify explicit evaluations but will have a weaker effect on implicit evaluations. Evidence for asymmetric effects of validity information on explicit and implicit evaluations may therefore be explained as resulting from expression-related processes, rather than the simultaneous operation of two independent learning processes.

Although validity information had an asymmetric effect on implicit and explicit evaluations in the long-delay condition, it is worth noting that it was still capable of partly qualifying implicit evaluations. Instead of reflecting the original valence of the behavioral descriptions, implicit evaluations did not differ between the two valence conditions when these descriptions were learned to be false. This attenuation deviates from Gregg et al.'s (2006) results, where validity information had no effect on implicit evaluations. Comparing the paradigms of the two sets of studies, there are at least two possible explanations for this difference. One explanation is that participants in the current experiments were forewarned that some information might turn out to be false, raising the possibility that our participants may have suspended belief in the observed contingencies until they knew their validity. To test this possibility, we conducted a replication of Experiment 3 in which participants were *not* informed, prior to the learning task, that some information might turn out to be false. The pattern of results was identical to that observed in Experiment 3, suggesting that the qualification of implicit evaluations observed in the current study is not the result of the suspension of belief during learning. A second possible explanation is that Gregg et al.'s (2006) experiments involved a much longer delay between learning of evaluative information and subsequent acquisition of validity information. Although in our study validity information was provided after all behavioral descriptions had been presented, participants in Gregg et al.'s studies completed measures of explicit and implicit evaluations before they were told that the initial behavioral information had been false. Thus, consistent with our emphasis on time as a critical factor, longer delays may allow consolidation of the initially formed associations. As a result, the impact of newly acquired validity information may decrease with increasing delays between the initial learning of evaluative information and the subsequent acquisition of validity information. Future research investigating the effects of continuously increasing intervals may help to clarify the role of time as a critical factor for the impact of validity information on implicit evaluations.

General Discussion

The present results provide converging evidence that during a single learning episode, in which the validity of the evaluation implied by an object-valence contingency can be quickly assessed, both explicit and implicit evaluations reflect a process of belief formation rather than distinct effects of belief-based and contingency-based learning processes. In Experiments 1 and 2, the perceived validity of behavioral descriptions of social targets qualified both explicit and implicit evaluations of these targets. This result suggests that validity information is incorporated into the mental representation of evaluative objects at the time of learning. Expanding on these findings, Experiment 3 demonstrated that validity information can have asymmetric influences later at the time of expression, when evaluative responses are reconsidered in light of additional information. Manipulating the delay between the presentation of contingencies and the presentation of validity information revealed that the impact of validity information was reduced for implicit, but not explicit, evaluations when validity information became available after a substantial delay.

Implications for Dissociations Between Explicit and Implicit Evaluations

Experimentally induced dissociations between explicit and implicit evaluations are often viewed as evidence for two independent learning mechanisms that may operate simultaneously on the basis of the same information (e.g., Gawronski & Bodenhausen, 2006; Rydell & McConnell, 2006; Strack & Deutsch, 2004). The experiments reported here suggest that the best interpretation of evaluative dissociations depends on the specific conditions of the learning situation. Although there is compelling evidence that dissociations can arise during learning when the valence of an object is manipulated using multiple, distinct sources of information (e.g., Rydell et al., 2006), the present experiments found no evidence for the *simultaneous* operation of dual learning processes on the basis of the *same* information. In these situations, when the evaluations implied by observed contingencies are immediately qualified by validity information, both explicit and implicit evaluations seem to be driven by a single process of belief formation.

Considering earlier evidence for evaluative dissociations arising during a single learning episode, it is worth noting that virtually all of this evidence can be straightforwardly interpreted as resulting from expression-related, rather than learning-related, processes. For example, using a learning procedure with a single, consciously available source of information, Ratliff and Nosek (2010) found that explicit evaluations showed the classic illusory correlation effect, whereas implicit evaluations reflected the actual contingencies of the observed information. Although they interpreted this finding as support for the independent operation of belief-based and contingency-based learning processes during a single learning episode, the dissociation can also be explained as a result of expression-related processes. In particular, illusory correlation effects may occur for explicit evaluations to the extent that infrequent information is more salient (Hamilton & Gifford, 1976) and salient information is given more weight in the process of generating an evaluative judgment. Importantly, such biases in the weighting of salient information may occur even if the relative strength of the underlying associations does not differ from the associations reflecting less salient information. From this perspective, the evaluative dissociation obtained by Ratliff and Nosek may not be the result of the simultaneous operation of two learning mechanisms but processes operating during the generation of evaluative judgments.

Implications for Dual-Process Theories

The present findings seem, prima facie, more compatible with recent arguments for single-process theories of learning, according to which all learning is the product of a single process of belief formation (e.g., Mitchell, De Houwer, & Lovibond, 2009). Evidence for dual-process learning in other situations notwithstanding, drawing such a conclusion in the present case seems premature for both empirical and epistemological reasons. First, it is always possible that in a single learning episode, the belief-based learning process is simply more powerful than the contingency-based learning process, thereby obscuring evidence for the operation of the latter process. Second, because theoretical entities cannot be observed directly, claims about their existence are not subject to direct empirical tests (Popper, 1934). Instead, existence claims have be evaluated indirectly by testing empirical predictions derived from assumptions about these entities. The underlying existence claims gain a measure of support when predictions are confirmed, but they will most likely be rejected when predictions repeatedly fail (Quine, 1969).

From this perspective, the current experiments can be understood as failing to confirm predictions derived from specific assumptions about contingency-based learning. The failure to confirm these predictions does not, however, conclusively demonstrate the nonexistence of contingency-based, as opposed to belief-based, learning processes. Indeed, there is evidence for the operation of both contingency-based and belief-based learning processes in certain conditions, such that specific conditions may promote the operation of one process and inhibit the operation of the other (e.g., Rydell et al., 2006). The current findings do, however, challenge the idea that two learning processes operate *simultaneously* and lead to divergent explicit and implicit evaluations on the basis of the *same* information (e.g., Gawronski & Bodenhausen, 2006). In these situations, it appears that conscious beliefs exert a strong qualifying influence on the evaluations implied by observed contingencies at the time of learning. The current findings thus impose a constraint on the conditions under which dual learning processes may produce divergent outcomes, though they do not rule out their existence.

Implications for Mental Control

The findings of the present research also shed light on the question raised at the outset of this article: Were Obama's supporters justified in worrying about an uncontrollable influence of negative images on voters? The answer seems to be that it depends. The experiments reported here suggest that beliefs about the invalidity of perceived object-valence contingencies (e.g., between *Obama* and *terrorism*) can qualify how that information is encoded and mentally represented, but only when the perceiver invalidates that information quickly. Otherwise, as the delay between the perception of a contingency and its invalidation increases, the mental association resulting from that contingency becomes more difficult to qualify (Gregg et al., 2006; Petty, Tormala, Briñol, & Jarvis, 2006). Thus, if a potential voter observed the illustration of Obama as a terrorist, she or he might be able to prevent the depicted link from being stored in memory by immediately rejecting it as false; failing to do so, however, might lead to the formation of a mental association that directly reflects it. This outcome would indeed be worrisome for Obama's supporters, as implicit evaluations have been shown to predict significant behavioral outcomes, including choice decisions in the political domain (e.g., Galdi et al., 2008; Payne et al., 2010). Nevertheless, there seems to be some room for control over the evaluations we form in typical learning situations, such that contingency-based learning may not be powerful enough to create mental associations that contradict our beliefs about what we observe. If we are puppets on a string to those who would seek to influence us, we at least have a brief opportunity to pull back.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interests with respect to the authorship and/or publication of this article.

Financial Disclosure/Funding

The authors disclosed receipt of the following financial support for the research and/or authorship of this article: This research was supported by the Ontario Graduate Scholarship Program and the Social Sciences and Humanities Research Council of Canada.

Note

1. The order of the two evaluation measures did not moderate the effect of valence and validity information on the measure of implicit evaluations, $F(1, 26) = 0.02$, $p = .879$, $\eta^2\ p < .01$, but did moderate the effect of valence and validity on the measure of explicit evaluations, $F(1, 26) = 6.59$, $p = .016$, $\eta^2_p = .20$. The order effect reflects a stronger effect of the Valence × Validity interaction on explicit evaluations when the measure of explicit evaluations was completed first, although in both cases the interaction remained significant. Specifically, when the measure of explicit evaluations was completed after the measure of implicit evaluations, the two-way interaction between valence and validity was relatively weaker, $F(1, 13) = 15.62$, $p = .002$, $\eta^2_p = .55$, compared to when it was completed first, $F(1, 13) = 133.60$, $p < .001$, $\eta^2_p = .91$. Because measurement order had no effect on the results in Experiments 2 and 3, we refrain from speculating on the nature of this effect.

References

Banaji, M. R. (2008, August 1). The science of satire: Cognition studies clash with "New Yorker" rationale. *Chronicle of Higher Education*, *54*, B13.

Bargh, J. A. (1994). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition. In R. S. Wyer, Jr. & T. K. Srull (Eds.), *Handbook of social cognition* (2nd ed., pp. 1-40). Hillsdale, NJ: Erlbaum.

Carraro, L., Gawronski, B., & Castelli, L. (2010). Losing on all fronts: The effects of negative versus positive person-based campaigns on implicit and explicit evaluations of political candidates. *British Journal of Social Psychology*, *49*, 453-470.

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.

Deutsch, R., Gawronski, B., & Strack, F. (2006). At the boundaries of automaticity: Negation as reflective operation. *Journal of Personality and Social Psychology*, *91*, 385-405.

Deutsch, R., Kordts-Freudinger, R., Gawronski, B., & Strack, F. (2009). Fast and fragile: A new look at the automaticity of negation processing. *Experimental Psychology*, *56*, 434-446.

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

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.

Galdi, S., Arcuri, L., & Gawronski, B. (2008). Automatic mental associations predict future choices of undecided decision-makers. *Science*, *321*, 1100-1102.

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., Cunningham, W. A., LeBel, E. P., & Deutsch, R. (2010). Attentional influences on affective priming: Does categorization influence spontaneous evaluations of multiply categorizable objects? *Cognition and Emotion*, *24*, 1008-1025.

Gawronski, B., & Walther, E. (2008). The TAR effect: When the ones who dislike become the ones who are disliked. *Personality and Social Psychology Bulletin*, *34*, 1276-1289.

Gawronski, B., Walther, E., & Blank, H. (2005). Cognitive consistency and the formation of interpersonal attitudes: Cognitive balance affects the encoding of social information. *Journal of Experimental Social Psychology*, *41*, 618-626.

Gregg, A. P., Seibt, B., & Banaji, M. R. (2006). Easier done than undone: Asymmetry in the malleability of implicit preferences. *Journal of Personality and Social Psychology*, *90*, 1-20.

Hamilton, D. L., & Gifford, R. K. (1976). Illusory correlation in interpersonal perception: A cognitive basis of stereotypic judgments. *Journal of Experimental Social Psychology*, *12*, 392-407.

Hofmann, W., Gschwendner, T., Nosek, B. A., & Schmitt, M. (2005). What moderates implicit-explicit consistency? *European Review of Social Psychology*, 16, 335-390.

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

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

Payne, B. K., Krosnick, J. A., Pasek, J., Lelkes, Y., Akhtar, O., & Tompson, T. (2010). Implicit and explicit prejudice in the 2008 American presidential election. *Journal of Experimental Social Psychology*, *46*, 367-374.

Petty, R. E., Tormala, Z. L., Briñol, P., & Jarvis, W. B. G. (2006). Implicit ambivalence from attitude change: An exploration of the PAST model. *Journal of Personality and Social Psychology*, *90*, 21-41.

Popper, K. R. (1934). *Logic of scientific discovery*. New York, NY: Basic Books.

Quine, W. V. O. (1969). Ontological relativity. In *Ontological relativity and other essays* (pp. 26-68). New York, NY: Columbia University Press.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, *114*, 510-532.

Ratliff, K. A., & Nosek, B. A. (2010). Creating distinct implicit and explicit attitudes with an illusory correlation paradigm. *Journal of Experimental Social Psychology*, *46*, 721-728.

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.

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

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

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: Theory*, *measurement*, *and applications* (pp. 95-116). New York, NY: Guilford.