COGNITION AND EMOTION
2009, 23 (6), 1118–1152

Psychology Press
Taylor & Francis Group

I like you, I like you not: Understanding the formation of context-dependent automatic attitudes

Robert J. Rydell
Indiana University, Bloomington, IN, USA

Bertram Gawronski
University of Western Ontario, London, Ontario, Canada

Previous research has shown that automatic evaluations can be highly context dependent. Expanding on past research demonstrating context effects for existing attitudes toward familiar objects, the present research examined basic principles that guide the *formation* of context-dependent versus context-independent automatic attitudes. Results from four experiments showed that: (a) newly formed attitudes generalised to novel contexts when prior experiences with the attitude object were evaluatively homogeneous; (b) when prior experiences were evaluatively heterogeneous, automatic evaluations became context sensitive, such that they reflected the contingency between the valence of prior experiences and the context in which these experiences occurred; and (c) when prior experiences were evaluatively heterogeneous across different contexts, novel contexts elicited automatic evaluations that reflected the valence of first experiences with the attitude object. Implications for research on automatic evaluation and attitude change are discussed.

Keywords: Attitude formation; Attitude change; Automatic evaluation; Context effects; Implicit measures.

One of the most important discoveries in attitude research during the last quarter century was the finding that attitudes can be activated automatically (Fazio, Sanbonmatsu, Powell, & Kardes, 1986). This finding led to the

Correspondence should be addressed to: Robert J. Rydell, Department of Psychological and Brain Sciences, Indiana University, 1101 E. 10th St., Bloomington, IN 47405, USA. E-mail: rjrydell@indiana.edu

This research was supported by PHS Grant MH 63762, Canada Research Chairs Program Grant 202555, Social Sciences and Humanities Research Council of Canada Grant 410-2005-1339, and Early Researcher Award ER07-03-067 from the Ontario Ministry for Research and Innovation.

We would like to thank the 18th and 19th Annual Duck Conference on Social Cognition for facilitating this research.

www.psypress.com/cogemotion DOI: 10.1080/02699930802355255

development of a vast number of indirect attitude measures (e.g., De Houwer, 2003; Fazio, Jackson, Dunton, & Williams, 1995; Greenwald, McGhee, & Schwartz, 1998; Nosek & Banaji, 2001; Payne, Cheng, Govorun, & Stewart, 2005), which were particularly designed to assess automatic evaluations of an attitude object (see Petty, Fazio, & Briñol, in press; Wittenbrink & Schwarz, 2007, for reviews).[1] Originally, automatic evaluations assessed by these measures were thought to reflect highly robust representations that had their roots in long-term socialisation experiences (e.g., Greenwald & Banaji, 1995; Rudman, 2004; Wilson, Lindsey, & Schooler, 2000; see Gawronski, LeBel, & Peters, 2007, for a discussion). However, this assumption has been challenged by an accumulating body of research showing that automatic evaluations can be highly context sensitive (see Blair, 2002; Gawronski & Bodenhausen, 2006, for reviews).

Even though context effects on automatic evaluations have been demonstrated for a large variety of objects and contexts, the particular conditions under which automatic evaluations are context dependent or context independent are not sufficiently well understood. Previous research has shown that automatic evaluations can differ as a function of social settings (e.g., Wittenbrink, Judd, & Park, 2001), social roles (e.g., Barden, Maddux, Petty, & Brewer, 2004), accessible exemplars (e.g., Dasgupta & Greenwald, 2001), salient categories (e.g., Mitchell, Nosek, & Banaji, 2003), goal pursuit (e.g., Ferguson & Bargh, 2004), and food deprivation (e.g., Seibt, Häfner, & Deutsch, 2007). These findings provide clear evidence that automatic evaluations can differ across contexts. However, none of these studies included conditions under which automatic evaluations turned out to be context independent. Hence, they remain silent about the particular factors that determine whether automatic evaluations are context dependent or context independent.

Needless to say, a clear understanding of the conditions under which automatic evaluations are context dependent or context independent is essential for accurate theorising about attitudes (see Fazio, 2007; Ferguson & Bargh, 2007; Gawronski & Bodenhausen, 2006; Schwarz, 2007). Thus, to overcome the limitations of previous research, the main goal of the present studies was to go beyond the mere demonstration of context effects on existing attitudes toward familiar objects, and to investigate principles guiding the *formation* of context-dependent and context-independent automatic attitudes (see Devine, 2001; Fazio & Olson, 2003). Using such an approach, we were particularly interested in (a) the causal role of experiential heterogeneity for context-dependent and context-independent

[1] In the present article, we use the term *automatic* in the sense of *unintentional* (i.e., spontaneous evaluative responses to an object that do not require an intention to evaluate that object).

automatic evaluations and (b) the principles associated with the contextualisation versus generalisation of automatic evaluations in novel contexts.

Experiential heterogeneity

The present research is based on the assumption that contextual influences on automatic evaluations depend on (a) the evaluative heterogeneity of an individual's associative representation of the attitude object, and (b) the contingency between the valence of available information and particular context cues during learning (Gawronski & Bodenhausen, 2007). Specifically, we expected that automatic evaluations of a given object tend to generalise across contexts when the associative representation of that object is evaluatively homogeneous. In this case, different sets of object-related input stimuli should activate evaluatively congruent patterns of associations in memory (see Bassili & Brown, 2005; Conrey & Smith, 2007) leading to similar evaluative responses across different contexts. However, when the associative representation of an object is evaluatively heterogeneous, automatic evaluations are assumed to become context sensitive. In this case, different sets of object-related input stimuli may activate different patterns of associations, leading to different evaluative responses as a function of the particular context (see Bassili & Brown, 2005; Conrey & Smith, 2007).

Obviously, a crucial factor that determines the evaluative heterogeneity of associative representations is the evaluative heterogeneity of prior experiences with the attitude object. In line with other researchers (e.g., Conrey & Smith, 2007; Gregg, Seibt, & Banaji, 2006; Petty, Tormala, Briñol, & Jarvis, 2006; Rydell & McConnell, 2006), we argue that new experiences do not erase old associations from memory, but simply add new associations to the already existing representation. Thus, whichever subset of these associations gets activated during future encounters depends on (a) the presence of relevant context cues when the object is encountered and (b) the contingency between these context cues and the valence of acquired information in the learning of that information (Bouton, 2005). Thus, drawing on the terminology of classical conditioning, one could argue that context cues serve as *occasion setters* that determine the nature of the conditioned response (i.e., automatic evaluation) that is elicited to a conditioned stimulus (i.e., attitude object) in the presence of these cues (see Schmajuk & Holland, 1998).

Notwithstanding the plausibility of these assumptions, the available evidence remains equivocal with regard to the proposed causal role of evaluatively heterogeneous representations. First, previous studies generally investigated context effects on existing attitudes toward familiar objects (e.g., Barden et al., 2004; Dasgupta & Greenwald, 2001; Mitchell et al., 2003;

Wittenbrink et al., 2001). Thus, even though it seems reasonable to assume that participants' experiences with these objects have been evaluatively heterogeneous (e.g., African Americans), participants' representations of these objects may have differed with regard to a number of aspects above and beyond the heterogeneity of evaluative experiences. Because we do not know the degree to which these aspects contributed to the obtained context dependency of automatic evaluations, previous research remains ambiguous as to whether heterogeneity of evaluative experiences is indeed *sufficient* to produce context-dependent automatic evaluations.

Second, even though previous research provided clear demonstrations of context-dependent automatic evaluations, these studies did not include conditions under which automatic evaluations generalised across contexts. Thus, given that participants' evaluative experiences with the attitude object were indeed heterogeneous, it seems possible that automatic attitudes are generally sensitive to the context in which they were formed, irrespective of whether prior evaluative experiences are homogeneous or heterogeneous. For instance, automatic attitudes may be activated only in contexts that correspond to the context in which the attitude was formed, thereby resulting in neutral evaluations in novel, unfamiliar contexts. From this perspective, previous research does not provide any evidence as to whether heterogeneity of evaluative experiences is indeed *necessary* to produce context-dependent automatic evaluations.

Third, many studies that have demonstrated context effects on automatic evaluations (e.g., Dasgupta & Greenwald, 2001; Mitchell et al., 2003; Wittenbrink et al., 2001) used context cues that themselves were either positive or negative (e.g., barbeque vs. ghetto). Thus, to the degree that these contexts elicit evaluative responses that are in line with their a priori valence (see Barden et al., 2004, for a notable exception), it remains ambiguous as to whether context cues simply provide additional information that contributes to an individual's overall automatic evaluative response or whether context cues indeed serve as *occasion setters* (see Schmajuk & Holland, 1998) for the activation of different responses to the attitude object. This question cannot be answered by investigating context effects on pre-existing attitudes and familiar contexts of positive and negative valence. Instead, it requires an experimental approach in which context-dependent automatic attitudes are created from scratch with context cues that do not have a strong a priori evaluative connotation.

Automatic evaluation in novel contexts

A second objective of the present research was to investigate how evaluatively heterogeneous experiences with an attitude object influence automatic evaluations in novel contexts. As stated above, we expected that automatic evaluations generalise to novel contexts when prior experiences with the object were evaluatively homogeneous. However, if the associative representation of an object is evaluatively heterogeneous—thereby evoking positive responses in one context and negative responses in another context—how is that object evaluated in novel contexts that are sufficiently dissimilar from the contexts in which evaluative information about the object has been acquired? From a general perspective, there are at least three possible outcomes. Specifically, automatic evaluations in novel contexts may reflect (a) the average or summation of all available information, (b) the valence of the information acquired first, or (c) the valence of the information acquired last.

First, automatic evaluations in novel contexts could reflect the average or summation of all available information (e.g., Anderson, 1971; Betsch, Kaufmann, Lindow, Plessner, & Hoffmann, 2006; Rydell & McConnell, 2006; Rydell, McConnell, Strain, Claypool, & Hugenberg, 2007), possibly leading to ambivalent evaluations when prior experiences have been evaluatively heterogeneous (e.g., de Liver, van der Pligt, & Wigboldus, 2007). Consistent with this assumption, research has shown that evaluatively heterogeneous experiences with an attitude object tend to result in neutral automatic evaluations, with automatic evaluations displaying features commonly obtained for ambivalent attitudes (e.g., Petty et al., 2006). However, deviating from the current emphasis on the contingency between context cues and evaluative experiences with an object, context cues in this research were held constant during the learning and measurement phases. As such, these studies do not provide conclusive evidence regarding automatic evaluations in novel contexts when prior experiences were evaluatively heterogeneous across different contexts.

Second, automatic evaluations in unfamiliar, neutral contexts could reflect the valence of earlier acquired information (e.g., Greenwald & Banaji, 1995; Rudman, 2004; Wilson et al., 2000). This assumption is consistent with research showing that earlier acquired information is more likely to influence automatic evaluations than subsequently acquired information (e.g., Gregg et al., 2006; Rudman, Phelan, & Heppen, 2007; Rydell & McConnell, 2006). For instance, Gregg et al. (2006) presented participants with evaluative information about two hypothetical groups, one of which was described in positive terms and the other in negative terms. After completing a measure of automatic evaluations, participants received additional information of the opposite valence, which was followed by the same measure of automatic evaluations. In line with Gregg et al.'s (2006) predictions, participants showed an automatic preference for the group that was described in positive terms in the first learning session regardless of whether automatic evaluations were assessed after the first or the second learning session. However, context cues in this research were also held constant during the learning and measurement phases. As such, these findings do not provide conclusive evidence regarding automatic evaluations in novel contexts when prior experiences were evaluatively heterogeneous across different contexts.

Finally, automatic evaluations in novel contexts may reflect the valence of the information acquired last. Even though we are not aware of any research showing recency effects on automatic evaluations, recency effects have been shown for self-reported, explicit evaluations (e.g., Miller & Campbell, 1959). Moreover, previous studies investigating changes in automatic evaluation as a function of newly acquired information (thus implying evaluatively heterogeneous representations) did not manipulate the contexts of information acquisition and automatic evaluation. Instead, context cues in these studies were typically held constant (e.g., Gregg et al., 2006; Petty et al., 2006; Rydell & McConnell, 2006). Hence, the available evidence does not allow any conclusions as to whether automatic evaluations in novel contexts reflect (a) the summation or average of all information, (b) the valence of information acquired first, or (c) the valence of information acquired last, when prior experiences were evaluatively heterogeneous across different contexts.

Contextualisation of ''exceptions''

Drawing on previous research on subtyping (e.g., Johnston & Hewstone, 1992; Weber & Crocker, 1983), we speculated that the second outcome, automatic evaluations reflecting the valence of earlier acquired information, might be the most likely one. Research in this area has shown that people tend to create subtypes when they receive information that violates stereotypical expectations about a social group. Importantly, the resulting subtypes seem to be stored as ''exceptions to the rule'', such that the original stereotype is still regarded as the accurate default case (e.g., Kunda & Oleson, 1995, 1997). From this perspective, it seems likely that context-specific, evaluatively heterogeneous experiences with an attitude object result in similar effects for automatic evaluations. Specifically, evaluative information that violates prior expectations may be stored as an exception to the rule, such that it influences automatic evaluations only in the particular context in which this information was acquired. Moreover, earlier acquired information may still function as the default case, thereby leading to a generalisation of earlier acquired attitudes to novel, unfamiliar contexts.

Similar conclusions can be drawn from research on context effects in the extinction of conditioned responses. Extinction refers to the phenomenon when the presentation of a conditioned stimulus (CS) devoid of the unconditioned stimulus (US) eliminates the conditioned response (CR; Rescorla & Wagner, 1972). However, in contrast to the assumption that

CS–US associations are completely eliminated during extinction, several studies by Bouton and colleagues suggest that extinction may involve the learning of new associations, resulting in two contextualised responses to the CS (e.g., Bouton, 2004; Bouton & Bolles, 1979; Bouton & King, 1983; Bouton & Schwartzentruber, 1986, 1989). Whereas the context during extinction serves to elicit the newly acquired (neutral) response, the original context still serves to reactivate the original, conditioned response. Importantly, the latter also seems true for novel, unfamiliar contexts, which have been shown to elicit the original, conditioned, response (e.g., Bouton & Bolles, 1979; Bouton & King, 1983; Bouton & Schwartzentruber, 1986, 1989; Harris, Jones, Bailey, & Westbrook, 2000). Such contextualisation of exceptions during extinction can also be applied to the notion of counter conditioning implied in the present research. Specifically, if an attitude object is initially associated with positive valence in one context (e.g., Tracey is associated with positivity at work) and subsequently with negative valence in a second context (e.g., Tracey is associated with negativity at home), it seems likely that the new, evaluatively inconsistent associations are activated only when the attitude object is encountered in the second context (i.e., Tracey at home). Moreover, in contexts that do not share features with any of the two contexts (e.g., Tracey at the gym), the original evaluation learned in the first context may serve as the default case, thereby implying a generalisation of earlier acquired evaluative information to novel, unfamiliar contexts.[2]

Overview of the present research

To provide deeper insights into the principles that guide the formation of context-dependent and context-independent automatic evaluations, we conducted four experiments. In line with current theorising (e.g., Gawronski & Bodenhausen, 2007), we predicted that automatic evaluations of an attitude object should be context independent (i.e., generalise across contexts) when participants' experiences with that object were evaluatively homogeneous. However, automatic evaluations of the same object should be context sensitive when participants' experiences with the attitude object were evaluatively heterogeneous. In the latter case, automatic evaluations in a given context should reflect the valence of the information that was associated with the attitude object in that context during learning. In addition, we were interested in how context-specific, evaluatively heterogeneous experiences would influence automatic evaluations in novel

[2] Note that such primacy effects of earlier acquired information in novel contexts may also occur when evaluatively inconsistent information is acquired in the same learning context and the novel context is sufficiently dissimilar to the context during learning (e.g., Bouton & Ricker, 1994). Even though this possibility is not explicitly addressed in the present studies, it seems an interesting question for future research.

contexts; namely, whether automatic evaluations in novel contexts reflect (a) the summation or average of all information, (b) the valence of information acquired first, or (c) the valence of information acquired last.

For this purpose, participants received information about a novel target person ("Bob") in a simple impression formation paradigm. This learning paradigm included a large amount of evaluative information about Bob in two different contexts. Afterwards, automatic evaluations of Bob were assessed in the two contexts of the learning task. Experiments 1 and 2 tested the prediction that heterogeneity of evaluative experiences determines the extent to which automatic evaluations are sensitive to contextual cues. Experiments 3 and 4 further examined how heterogeneity of evaluative experiences influences automatic evaluations in novel contexts. Toward that end, Experiments 3 and 4 used a paradigm similar to the one employed in Experiments 1 and 2, additionally including a novel context in the measurement of automatic evaluations.

EXPERIMENT 1

In an initial test of our predictions, participants learned about a visually presented target individual named Bob over the course of two learning sessions. In the first session, participants were presented with positive (negative) information about Bob in one of two different, evaluatively neutral contexts (i.e., a coloured background). In the second session, participants received evaluative information of the opposite valence in a different context (i.e., a differently coloured background). After each of the two learning sessions, participants completed a measure of automatic evaluation, in which a picture of Bob was presented within each of the two backgrounds. To test whether context effects were uniquely related to the specific target individual, the measure additionally included pictures of several unknown individuals against each of the two backgrounds. Based on the aforementioned theoretical considerations, we expected to obtain three sets of findings.

First, we expected automatic evaluations after the first learning session to reflect the valence of the presented information. Importantly, this effect should be context independent, such that it should emerge irrespective of the context in which Bob is presented during the measurement of automatic evaluations. This prediction implies a generalisation of automatic attitudes to novel contexts when prior experiences are evaluatively homogeneous. As mentioned earlier, such a generalisation stands in contrast to the possibility of ubiquitous contextualisation (e.g., Schwarz, 2007), implying that automatic attitudes are activated only in contexts that correspond to the context in which the attitude was formed.

Second, after the presentation of evaluatively opposite information about Bob in a different context, automatic evaluations should become context dependent, reflecting the valence of the information that was associated with a particular context. More precisely, automatic evaluations should still reflect the initial valence when Bob is encountered in the context of the first learning session; however, automatic evaluations should reflect the valence of the newly acquired information when Bob is encountered in the context of the second learning session.

Third, these context effects should be driven by the heterogeneous representation of the target individual (Ferguson & Bargh, 2007). As such, context effects should only emerge for the target individual, but they should not generalise to other unknown individuals that are evaluated in the same contexts.

Method

Participants and design. Eighty-seven undergraduates at the University of California, Santa Barbara, participated for research credit. The experiment consisted of a 2 (Order of Valence: positive first vs. negative first) $\times$ 2 (Valence–Context Match: positive–yellow, negative–blue vs. positive–blue, negative–yellow) $\times$ 2 (Context During Measurement: yellow vs. blue) $\times$ 2 (Time of Measurement: time 1 vs. time 2) mixed-model design, with the first two variables as between-subjects factors and that latter two as within-subjects factors. Data from three participants who were familiar with the meaning of the Chinese characters employed in the dependent measure (see below) were excluded from analyses.

Evaluative learning paradigm. To manipulate automatic evaluations, we utilised a modified version of Rydell and McConnell's (2006) evaluative learning paradigm. Participants were asked to form an impression of a target person named Bob based on written information about this person. Over the course of 100 trials, participants read about behaviours that Bob had performed while a picture of Bob was concurrently presented on the screen. Statement–picture pairs were presented for 5000 ms on the computer screen with an inter-trial interval of 1000 ms. During the first 50 trials, half of the participants were shown 40 positive behaviours and 10 neutral behaviours. The other half were presented with 40 negative behaviours and 10 neutral behaviours. During the second 50 trials, the valence of the behaviours was switched, such that participants who were initially presented with positive (negative) behaviours now received negative (positive) behaviours. The context in the evaluative learning task was manipulated by the background colour of the computer screen. For half of the participants, the behavioural information during the first 50 trials was presented on a blue background, and the behavioural information during the second 50 trials on a yellow background. The remaining half read the first 50 behaviours on a yellow background and the second 50 behaviours on a blue background.

Measure of automatic evaluations. As a measure of automatic evaluations, we used Payne et al.'s (2005) affect misattribution procedure (AMP). On each trial, participants were first presented with a face prime for 75 ms. The face was then replaced by a blank screen for 125 ms, which was followed by a Chinese character for 100 ms. Immediately after the presentation of the Chinese character, a black-and-white-pattern mask was presented, and participants were asked to indicate whether they considered the Chinese character as more pleasant or less pleasant than the average Chinese character. Specifically, participants were asked to press a right-hand key (*Numpad 5*) when they considered the Chinese character as more pleasant than average and a left-hand key (*A*) when they considered the Chinese character as less pleasant than average. Following the procedures of Payne et al. (2005), participants were told that the face primes tend to influence evaluations of the Chinese characters, and that they should do their absolute best not to let the faces bias their judgments of the Chinese characters. The AMP consisted of 100 trials. Half of the trials used the picture of Bob as prime stimulus; the remaining half used pictures of four unknown individuals as primes. In addition, we manipulated the background colour during the 75 ms presentation of the face primes, with half of the prime stimuli being presented on a yellow background and the remaining half being presented on a blue background. Order of trials was determined randomly by the computer. Automatic evaluations were assessed twice using the same measure, once after the first 50 trials of the evaluative learning task (time 1) and once after the second set of 50 trials (time 2).

Results

To examine context effects on automatic evaluations, we calculated the proportion of *more pleasant* responses for each of the four types of primes (i.e., Bob–yellow; Bob–blue; unknown–yellow; unknown–blue) at time 1 and time 2, respectively. Data were then collapsed across the counterbalanced background colours to reflect the two major within-subjects factors of Context Valence (i.e., positive vs. negative) and Time of Measurement (i.e., time 1 vs. time 2), and the between-subjects factor of Valence Order (i.e., positive first vs. negative first).

Submitted to a 2 (Valence Order: positive first vs. negative first) $\times$ 2 (Context Valence: positive vs. negative) $\times$ 2 (Time of Measurement: time 1 vs. time 2) mixed-model ANOVA, automatic evaluations of Bob revealed a significant main effect of Context Valence, $F(1,\ 82) = 22.83$, $p < .001$,

$\eta^2 = .218$, a significant two-way interaction of Valence Order and Time, $F(1, 82) = 12.92$, $p = .01$, $\eta^2 = .136$, a significant two-way interaction of Context Valence and Time, $F(1, 82) = 20.00$, $p < .001$, $\eta^2 = .196$, and, more important, a significant three-way interaction of Context Valence, Valence Order, and Time, $F(1, 82) = 4.10$, $p < .05$, $\eta^2 = .048$ (see Figure 1). To specify this interaction, we conducted separate 2 (Context Valence) × 2 (Time of Measurement) ANOVAs for each of the two valence order conditions.

When positive information was presented first (see Figure 1, left panel), the ANOVA revealed a significant main effect of Context Valence, $F(1, 39) = 16.67$, $p < .001$, $\eta^2 = .299$, a significant main effect of Time, $F(1, 39) = 5.92$, $p = .02$, $\eta^2 = .132$, and, more important, a significant two-way interaction of Context Valence and Time, $F(1, 39) = 15.12$, $p < .001$, $\eta^2 = .279$. Consistent with our predictions, automatic evaluations of Bob at time 1 were highly positive irrespective of context valence, $F(1, 39) = 0.11$, $p = .74$, $\eta^2 = .003$. In contrast, automatic evaluations at time 2 were more positive in the positive context as compared to the negative context, $F(1, 39) = 23.75$, $p < .001$, $\eta^2 = .378$. Moreover, automatic evaluations became less positive as a function of time in the negative context, $F(1, 39) = 14.52$, $p < .001$, $\eta^2 = .270$, but not in the positive context, $F(1, 39) = 1.98$, $p = .17$, $\eta^2 = .048$.

When negative information was presented first (see Figure 1, right panel), the ANOVA also revealed a significant main effect of Context Valence, $F(1, 43) = 6.53$, $p = .01$, $\eta^2 = .132$, a significant main effect of Time, $F(1, 43) = 7.05$, $p = .01$, $\eta^2 = .141$, and, more important, a significant two-way interaction, $F(1, 43) = 4.50$, $p < .05$, $\eta^2 = .095$. Mirroring the pattern obtained in the positive first condition, automatic evaluations of Bob at time 1 were highly negative irrespective of Context Valence, $F(1, 43) = 0.41$,

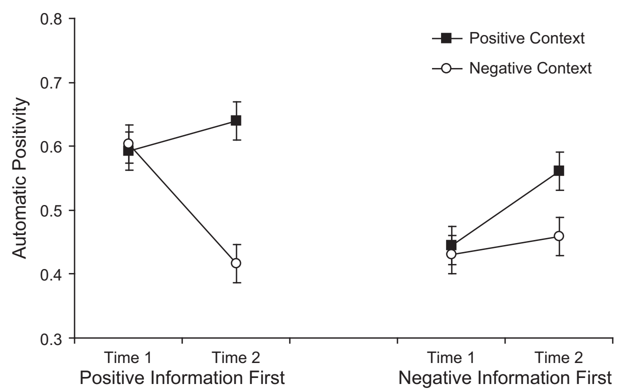

Figure 1. Automatic positivity toward the target individual as a function of context valence (positive vs. negative), information order (positive first vs. negative first), and time (time 1 vs. time 2), Experiment 1.

$p = .53$, $\eta^2 = .009$. In contrast, automatic evaluations at time 2 were more negative in the negative context as compared to the positive context, $F(1, 43) = 7.27$, $p = .01$, $\eta^2 = .145$. Moreover, automatic evaluations became more positive as a function of time in the positive context, $F(1, 43) = 8.76$, $p = .005$, $\eta^2 = .169$, but not in the negative context, $F(1, 43) = 1.04$, $p = .31$, $\eta^2 = .024$.

To further illustrate the predicted differences in context sensitivity, we also conducted separate 2 (Valence Order) × 2 (Context Valence) ANOVAs for the two time conditions (i.e., time 1 vs. time 2). For automatic evaluations assessed after the first learning session, this ANOVA only revealed a highly significant main effect of Valence Order, $F(1, 82) = 16.55$, $p < .001$, $\eta^2 = .168$, indicating that automatic evaluations of Bob were more positive when participants received positive information about Bob than when they received negative information. Most important, there was neither a significant main effect of the context in which Bob was presented, $F(1, 82) = 0.01$, $p = .94$, $\eta^2 < .001$, nor a significant interaction, $F(1, 82) = 0.41$, $p = .52$, $\eta^2 = .005$, supporting our prediction of context-independent evaluations when prior experiences were evaluatively homogeneous. However, for automatic evaluations assessed after the second learning session, the same ANOVA revealed a highly significant main effect of Context Valence, $F(1, 82) = 30.38$, $p < .001$, $\eta^2 = .270$, indicating that Bob was evaluated more positively when he was presented in a context associated with positive information than when he was presented in a context associated with negative information. This effect was qualified by a significant two-way interaction, $F(1, 82) = 4.23$, $p = .04$, $\eta^2 = .049$, suggesting that context-dependent differences in automatic evaluations at time 2 were somewhat more pronounced when positive information was presented first, $F(1, 39) = 23.75$, $p < .001$, $\eta^2 = .378$, than when negative information was presented first, $F(1, 43) = 7.27$, $p = .01$, $\eta^2 = .145$, though context effects reached statistical significance in both order conditions.

To test whether the obtained context effects were uniquely related to the target individual for which the evaluative information was presented, automatic evaluations of unknown individuals were submitted to the same 2 (Valence Order) × 2 (Context Valence) × 2 (Time of Measurement) mixed-model ANOVA. No main or interaction effect reached statistical significance (all $Fs < 2.38$, all $ps > .13$). The crucial three-way interaction of Valence Order, Context Valence, and Time was far from statistical significance, $F(1, 82) = 0.52$, $p = .47$, $\eta^2 = .006$.

Discussion

The results from Experiment 1 are consistent with our contention that context effects on automatic evaluations are driven by the activation of different subsets of information associated with the attitude object across different contexts. When attitudes were initially formed in one context, the valence of the acquired information determined automatic evaluations irrespective of the context in which automatic evaluations were measured. This result supports the assumption that automatic evaluations generalise to novel contexts when prior experiences with the attitude object were evaluatively homogeneous. However, the present finding is inconsistent with the assumption that automatic evaluations are ubiquitously context sensitive (e.g., Schwarz, 2007), such that automatic attitudes are activated only in contexts that correspond to the context in which the attitude was formed. Nevertheless, automatic evaluations became highly context sensitive after evaluative information of the opposite valence was learned in a different context, such that automatic evaluations reflected the valence of the information that was associated with the attitude object in a particular context. In other words, the emergence of context effects depended on (a) the heterogeneity of evaluative experiences with the attitude object, and (b) the contingency between context cues and particular subsets of evaluative information.

EXPERIMENT 2

Even though the results from Experiment 1 are consistent with our predictions, there are three potential problems that limit the conclusions that can be drawn from this experiment. First, utilising two separate learning sessions, comprised exclusively of positive or negative information, could artificially promote the development of associative clusters including the attitude object, context cues, and evaluative information. In an extreme variant of this argument, one could claim that Bob in the first learning session may have been perceived as a different person than Bob in the second learning session. Even though we consider an identification of the target person Bob as two different individuals as rather unlikely (the two learning session used the same picture of Bob and simply varied the colour of the background), Experiment 2 included a single learning session in which positive and negative information was randomly interspersed.

Another possible concern regarding Experiment 1 is that the employed pre–post design confounded heterogeneity of evaluative experiences and amount of information. Specifically, participants had learned twice as many behaviours at time 2 (i.e., 100 behaviours) compared to time 1 (i.e., 50 behaviours). Thus, it seems possible that the overall amount of evaluative experience, rather than evaluative heterogeneity of that experience, determined the context dependency of automatic evaluations. Even though this account cannot explain the obtained pattern of automatic evaluations without making reference to the specific nature of evaluative experiences, Experiment 2 aimed to address this concern by manipulating experiential heterogeneity independent of the overall amount of attitudinal experiences.

Finally, one could object to Experiment 1's utilisation of a recently introduced measure of automatic evaluation—Payne et al.'s (2005) affect misattribution procedure (AMP)—which has not yet received much empirical support aside from Payne et al.'s original studies (but see Payne, Govorun, & Arbuckle, 2008; Payne, McClernon, & Dobbins, 2007). In particular, the AMP has not yet been shown to be interchangeable with more established measures of automatic evaluation. In fact, recent research suggests that measures of automatic evaluation differ in important ways, which can lead to different outcomes of the same experimental manipulation for different measures (e.g., Deutsch & Gawronski, 2008; Gawronski & Bodenhausen, 2005; Gawronski, Cunningham, LeBel, & Deutsch, 2008). For this reason, Experiment 2 sought to rectify any problems regarding the interpretation of the AMP data by additionally including a variant of Fazio et al.'s (1995) affective priming task.

In summary, Experiment 2 used an evaluative learning paradigm similar to the one employed in Experiment 1. However, to address the aforementioned methodological issues, this paradigm included a few modifications. To resolve the confounding of heterogeneity of evaluative experiences and amount of information, all participants were presented with a total of 80 behaviours. Half of the behaviours were presented on a blue background, the remaining half were presented on a yellow background. Valence of the behaviours was manipulated independently for the two backgrounds, such that behaviours of each background were either positive or negative. Depending on the particular matching between valence and background, this manipulation implies either evaluatively homogeneous experiences (i.e., positive–positive; negative–negative) or evaluatively heterogeneous experiences (i.e., positive–negative; negative–positive). To address the concern that context effects may depend on artificially induced clustering processes, the presentation of evaluative information was interspersed randomly throughout the learning session. Finally, to demonstrate the generality of the obtained pattern of results across measures, participants completed two different measures of automatic evaluation: Payne et al.'s (2005) affect misattribution procedure and Fazio et al.'s (1995) affective priming task.

Method

Participants and design. Sixty-one undergraduates at the University of California, Santa Barbara, participated for research credit. The experiment consisted of a 2 (Yellow-Valence: positive vs. negative) × 2 (Blue-Valence: positive vs. negative) × 2 (Context During Measurement: yellow vs. blue) mixed-model design, with the first two variables as between-subjects factors and the latter as a within-subjects factor.

Evaluative learning paradigm. To manipulate automatic evaluations, we utilised a modified version of the evaluative learning paradigm employed in Experiment 1. Participants were asked to form an impression of a target person named Bob based on written information about that person. Over the course of 80 trials, participants read about behaviours that Bob had performed while his picture was concurrently presented on the screen. As with Experiment 1, statement–picture pairs were presented for 5000 ms on the computer screen with an inter-trial interval of 1000 ms. Half of the behaviours were presented on a yellow screen; the remaining half were presented on a blue screen. The valence of the presented behaviours was manipulated independently for each of the two backgrounds, such that behaviours of each background were either consistently positive or consistently negative. This manipulation resulted in four between-subjects conditions, reflecting the matching of background colour and valence. Thus, the valence of the information about Bob was consistent across contexts in two of the four conditions (i.e., yellow–positive/blue–positive; yellow–negative/blue–negative) and inconsistent in the other two conditions (i.e., yellow–positive/blue–negative; yellow–negative/blue–positive). Participants were presented with all 80 behaviours about Bob in a single learning session. The order of the trails was determined by the computer, with background colour and valence being randomly interspersed throughout the learning session.

Measures of automatic evaluations. As a measure of automatic evaluation, we again employed Payne et al.'s (2005) AMP. The procedure of the AMP was identical to the one employed in Experiment 1. In addition, all participants completed a variant of Fazio et al.'s (1995) affective priming task. On each trial of the affective priming task, participants were presented with a fixation cross in the centre of the screen for 1000 ms, which was followed by one of the primes used in the AMP (i.e., Bob on a blue screen, Bob on a yellow screen, an unknown person on a blue screen, and unknown person on a yellow screen). After 200 ms, the prime stimulus was replaced by a positive or negative target word. Participants' task was to indicate the valence of the target word as quickly as possible. Specifically, participants were asked to press a right-hand key (*Numpad 5*) when they saw one of twenty positive words (e.g., *paradise*, *summer*) and a left-hand key (*A*) when they saw one of twenty negative words (e.g., *evil*, *sickness*). Each target word was presented once with each of the four prime types, resulting in a total of 160 trials. Order of trials was randomly determined by the computer. Participants were required to respond within 1000 ms after the onset of the target word. If participants did not respond within 1000 ms, the message *Please try to respond faster!* was displayed for 1500 ms in the centre of the screen (see Klauer & Teige-Mocigemba, 2007). In addition, incorrect responses were indicated with the message *Error!* appearing for 1000 ms in the centre of the screen.

Results

Affect misattribution procedure. To examine context effects on automatic evaluations as measured by the AMP, we calculated the proportion of *more pleasant* responses for each of the four types of primes (i.e., Bob–yellow; Bob–blue; unknown–yellow; unknown–blue), respectively. Data were then aggregated as a function of whether the evaluative information about Bob was consistent or inconsistent across the two contexts. Specifically, we set the valence associated with the yellow background as a standard and then coded whether the valence associated with the blue background was consistent or inconsistent with that valence.

Submitted to a 2 (Yellow-Valence: positive vs. negative) × 2 (Consistency: blue consistent with yellow vs. blue inconsistent with yellow) × 2 (Measurement Context: blue vs. yellow) mixed-model ANOVA, automatic evaluations of Bob revealed a significant main effect of Yellow-Valence, $F(1, 57) = 8.27$, $p = .006$, $\eta^2 = .127$, a significant two-way interaction of Yellow-Valence and Consistency, $F(1, 57) = 10.05$, $p = .002$, $\eta^2 = .136$, a significant two-way interaction of Yellow-Valence and Measurement Context, $F(1, 57) = 18.46$, $p < .001$, $\eta^2 = .245$, a significant two-way interaction of Consistency and Measurement Context, $F(1, 57) = 4.43$, $p = .04$, $\eta^2 = .072$, and, more important, a significant three-way interaction of Yellow-Valence, Consistency, and Measurement Context, $F(1, 57) = 19.69$, $p < .001$, $\eta^2 = .257$ (see Figure 2). To specify this interaction in terms of our hypotheses, we conducted separate 2 (Yellow-Valence) × 2 (Measurement Context) ANOVAs for each of the two consistency conditions.

When information was consistent across the two contexts (see Figure 2, left panel), the ANOVA revealed a significant main effect of Yellow-Valence, $F(1, 30) = 24.14$, $p < .001$, $\eta^2 = .446$. Consistent with our predictions, automatic evaluations were more positive when the yellow context (and thus the blue context as well) was positive than when the yellow context (and thus the blue context as well) was negative. In addition, a marginally significant main effect of measurement context indicated a tendency toward more positive evaluations when measured in the blue context as opposed to the yellow context, $F(1, 30) = 3.62$, $p = .07$, $\eta^2 = .108$. The interaction of Yellow-Valence and Measurement Context was far from statistical significance, $F(1, 30) = 0.01$, $p = .91$, $\eta^2 < .001$.

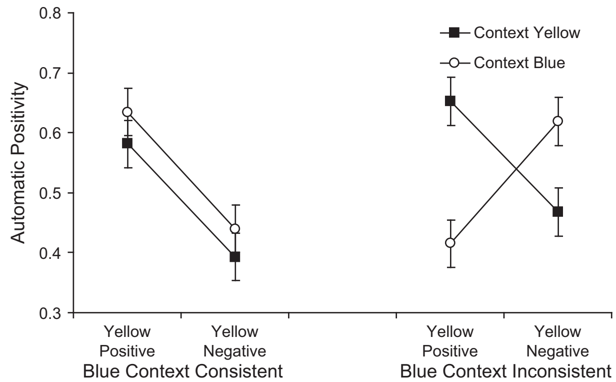

Figure 2. Automatic positivity toward the target individual on the affect misattribution procedure as a function of measurement context (blue vs. yellow), consistency (blue-valence consistent with yellow-valence vs. blue-valence inconsistent with yellow-valence), and yellow-valence (positive vs. negative), Experiment 2.

When information was inconsistent across the two contexts (see Figure 2, right panel), the same ANOVA revealed a highly significant two-way interaction of Yellow-Valence and Measurement Context, $F(1, 27) = 28.19$, $p < .001$, $\eta^2 = .511$. When yellow-valence was positive (and thus blue-valence was negative), automatic evaluations were more positive when measured in the yellow context than in the blue context, $F(1, 15) = 18.68$, $p < .001$, $\eta^2 = .555$. In contrast, when yellow-valence was negative (and thus blue-valence was positive), automatic evaluations were more positive when measured in the blue context than in the yellow context, $F(1, 12) = 11.41$, $p = .005$, $\eta^2 = .487$. Moreover, automatic evaluations measured in the yellow context were more positive when yellow-valence was positive than when it was negative, $F(1, 27) = 7.57$, $p = .01$, $\eta^2 = .219$. Conversely, automatic evaluations measured in the blue context were more positive when yellow-valence was negative than when yellow-valence was positive, $F(1, 27) = 12.43$, $p = .002$, $\eta^2 = .315$.

To test whether the obtained context effects were uniquely related to the target individual Bob, automatic evaluations of unknown individuals were submitted to the same 2 (Yellow-Valence) × 2 (Consistency) × 2 (Measurement Context) ANOVA. No main or interaction effect reached statistical significance (all Fs < 1.67, all ps $> .20$). The crucial three-way interaction between Yellow-Valence, Consistency, and Measurement Context was far from statistical significance, $F(1, 57) = 0.67$, $p = .42$, $\eta^2 = .012$.

Affective priming task. To corroborate the validity of the findings obtained with Payne et al.'s (2005) AMP, we additionally analysed context effects on automatic evaluations as measured by Fazio et al.'s (1995) affective priming task. For this purpose, we first eliminated response latencies from incorrect responses and responses that exceeded the response deadline of 1000 ms (6.2%). We then calculated scores of automatic positivity by subtracting the mean response latency for the identification of positive words from the mean response latency for the identification of negative words given a particular prime category (i.e., Bob–yellow; Bob–blue; unknown–yellow; unknown–blue).[3] As with the procedures described for the AMP, data were then aggregated as a function of whether the information about Bob was consistent or inconsistent across the two different contexts.

Submitted to a 2 (Yellow-Valence: positive vs. negative) $\times$ 2 (Consistency: blue consistent with yellow vs. blue inconsistent with yellow) $\times$ 2 (Measurement Context: blue vs. yellow) mixed-model ANOVA, automatic evaluations of Bob revealed a significant two-way interaction of Yellow-Valence and Measurement Context, $F(1, 57) = 5.17$, $p = .03$, $\eta^2 = .083$, and, more important, a significant three-way interaction of Yellow-Valence, Consistency, and Measurement Context, $F(1, 57) = 15.95$, $p < .001$, $\eta^2 = 219$ (see Figure 3). To specify this interaction in terms of the present hypothesis, we conducted separate 2 (Yellow-Valence) $\times$ 2 (Measurement Context) ANOVAs for each of the two consistency conditions.

When information was consistent across the two different contexts (see Figure 3, left panel), the ANOVA revealed a marginally significant main effect of Yellow-Context, $F(1, 30) = 2.98$, $p = .09$, $\eta^2 = .090$. Specifically, automatic evaluations of Bob tended to be more positive when the yellow context (and thus the blue context as well) was positive than when the yellow context (and thus the blue context as well) was negative. The interaction of Yellow-Valence and Measurement Context failed to reach statistical significance, $F(1, 30) = 1.26$, $p = .27$, $\eta^2 = .040$.

When information was inconsistent across the two different contexts (see Figure 3, right panel), the same ANOVA revealed a highly significant two-way interaction between Yellow-Valence and Measurement Context, $F(1, 27) = 25.22$, $p < .001$, $\eta^2 = .483$. When yellow-valence was positive (and thus blue-valence was negative), automatic evaluations were more positive when measured in the yellow context than in the blue context, $F(1, 15) = 20.36$, $p < .001$, $\eta^2 = .576$. In contrast, when yellow-valence was negative (and thus blue-valence was positive), automatic evaluations were more positive when

[3] Note that responses to positive target words are typically faster than responses to negative words, thereby promoting scores higher than zero for the present scoring. Thus, the resulting priming scores should *not* be interpreted in an absolute manner, such that scores higher than zero would indicate a positive response and scores lower than zero would indicate a negative response. Instead, priming scores should only be interpreted in a *relative* manner, such that higher scores indicate more positive responses.

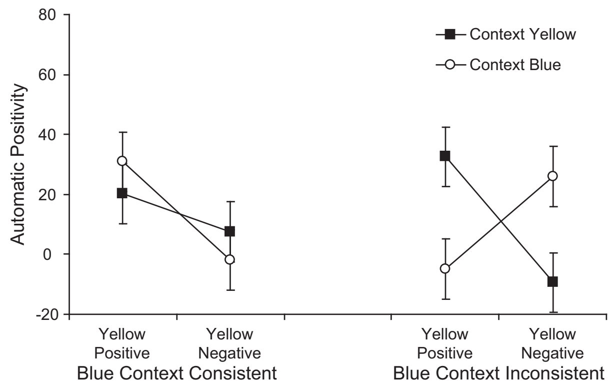

Figure 3. Automatic positivity toward the target individual on the affective priming task as a function of measurement context (blue vs. yellow), consistency (blue-valence consistent with yellow-valence vs. blue-valence inconsistent with yellow-valence), and yellow-valence (positive vs. negative), Experiment 2.

measured in the blue context than in the yellow context, $F(1, 12) = 8.07$, $p = .02$, $\eta^2 = .402$. Moreover, automatic evaluations measured in the yellow context were more positive when yellow-valence was positive than when it was negative, $F(1, 27) = 22.91$, $p < .001$, $\eta^2 = .459$. Conversely, automatic evaluations measured in the blue context were more positive when yellow-valence was negative than when yellow-valence was positive, $F(1, 27) = 5.61$, $p = .03$, $\eta^2 = .172$.

To test whether the obtained context effects were uniquely related to the target individual Bob, automatic evaluations of unknown individuals were submitted to the same 2 (Yellow-Valence) × 2 (Consistency) × 2 (Measurement Context) ANOVA. No main or interaction effect reached statistical significance (all Fs < 2.49, all ps $> .12$). The crucial three-way interaction between Yellow-Valence, Consistency, and Measurement Context failed to reach statistical significance, $F(1, 57) = 1.43$, $p = .24$, $\eta^2 = .024$.

Discussion

Experiment 2 provided further support for our assumption that context effects on automatic evaluations are driven by the activation of different subsets of information associated with the attitude object across different contexts. In the present study, automatic evaluations were consistent across contexts when experiences with the attitude object were homogeneous. However, automatic evaluations were highly context sensitive when experiences with the attitude object were heterogeneous. In addition, Experiment 2 rules out several alternative explanations for the results of Experiment 1. First, heterogeneity of evaluative experiences was manipulated independent of the overall amount of information, thereby ruling out alternative explanations in terms of amount of information. Second, the learning paradigm employed in Experiment 2 included an interspersed presentation of positive and negative information in different contexts, thereby ruling out alternative explanations in terms of artificially induced associative clustering. Finally, Experiment 2 showed identical patterns of results for two different measures of automatic evaluation—Payne et al.'s (2005) affect misattribution procedure and Fazio et al.'s (1995) affective priming task—thereby providing clear evidence for the reliability of the obtained effects across measures.

Given the obtained context sensitivity of automatic evaluations when experiences with the attitude object were evaluatively heterogeneous, how people respond if that object is encountered in a novel, unfamiliar context is an open question. As outlined in the introduction, there are at least three possibilities of how experiential heterogeneity may influence automatic evaluations in novel contexts: (a) automatic evaluations may reflect the average or sum of all available information; (b) automatic evaluations may reflect the valence of the information acquired first; or (c) automatic evaluations may reflect the valence of the information acquired last. These possibilities were addressed in the final two experiments.

EXPERIMENT 3

To address the question of how heterogeneity of evaluative experiences influences automatic evaluations in novel contexts, Experiment 3 used a two-session learning procedure similar to the one employed in Experiment 1. In the first learning session, participants were presented with negative information about Bob in one context. The second session included positive information presented in a different context. After each of the two learning sessions, participants completed a measure of automatic evaluation, in which Bob was present against either (a) the first (negative) context, or (b) the second (positive) context, or (c) a novel (neutral) context. Drawing on previous evidence on subtyping (e.g., Johnston & Hewstone, 1992; Weber & Crocker, 1983) and contextualisation in the extinction of conditioned responses (e.g., Bouton & Bolles, 1979; Bouton & King, 1983; Bouton & Schwartzentruber, 1986, 1989; Harris et al., 2000), we expected that new information that contradicts prior expectancies may be stored as an "exception to the rule", such that the initial expectation is still regarded as an accurate, default case. Hence, evaluative information that is inconsistent with an earlier acquired attitude may determine subsequent automatic evaluations only in the context in which this information was acquired, with the earlier acquired attitude still generalising to novel contexts.

Method

Participants and design. Forty-four undergraduates at the University of California, Santa Barbara, participated for research credit. The experiment consisted of a 2 (Order of Background: yellow first, blue second vs. blue first, yellow second) $\times$ 3 (Context During Measurement: yellow vs. white vs. blue) $\times$ 2 (Time of Measurement: time 1 vs. time 2) mixed-model design, with the first factor varying between-subjects and the latter two varying within-subjects.

Procedure. The evaluative learning paradigm was identical to the one employed in Experiment 1. As a measure of automatic evaluations, we again used Payne et al.'s (2005) affect misattribution procedure (AMP). The procedure was largely identical to the AMP employed in Experiment 1, the only differences being: (a) a reduction of the overall number of trials from 100 to 90, and (b) the inclusion of a white background to represent a novel, neutral background for one third of the trials. Thus, each of the two face primes (i.e., Bob, unknown) was presented 15 times on each of the three backgrounds (i.e., yellow, blue, white). As in Experiment 1, automatic evaluations were assessed twice using the same measure, once after the first session of the evaluative learning task (time 1) and once after the second session (time 2).

Results

To examine context effects on automatic evaluations, we first calculated the proportion of *more pleasant* responses for each of the three types of prime categories (i.e., Bob–yellow; Bob–blue; Bob–white) at time 1 and time 2, respectively. Data were then aggregated as a function of context valence (i.e., positive vs. neutral vs. negative), collapsing across the method-factor of background colour.

Submitted to a 3 (Context Valence: positive vs. neutral vs. negative) $\times$ 2 (Time of Measurement: time 1 vs. time 2) within-subjects ANOVA, automatic evaluations of Bob revealed a significant two-way interaction, $F(2, 86) = 8.21$, $p = .001$, $\eta^2 = .160$ (see Figure 4). Replicating the pattern obtained in Experiment 1, automatic evaluations of Bob at time 1 did not differ as a function of context valence, $F(2, 86) = 1.61$, $p = .21$, $\eta^2 = .036$. In contrast, automatic evaluations at time 2 showed a significant effect of the context, $F(2, 86) = 3.85$, $p = .03$, $\eta^2 = .082$. Again replicating the pattern obtained in Experiment 1, automatic evaluations were more positive in the positive context than in the negative context, $F(1, 43) = 6.82$, $p = .01$, $\eta^2 = .137$. More important for the present question, automatic evaluations in the neutral context were less positive than automatic evaluations in the

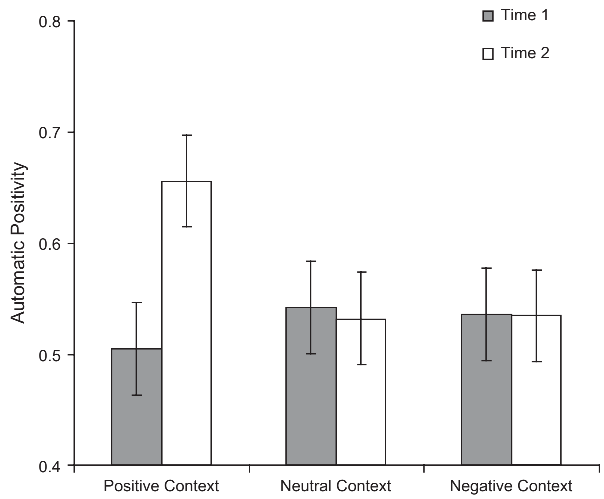

Figure 4. Automatic positivity toward the target individual as a function of context valence (positive vs. neutral vs. negative) and time (time 1 vs. time 2) when negative information was presented at time 1 and positive information at time 2, Experiment 3.

positive context, $F(1, 43) = 4.16$, $p < .05$, $\eta^2 = .088$, but did not differ from automatic evaluations in the negative context, $F(1, 43) = 0.29$, $p = .87$, $\eta^2 = .001$. Moreover, automatic evaluations became more positive as a function of time in the positive context, $F(1, 43) = 8.74$, $p = .01$, $\eta^2 = .165$, but not in the neutral context, $F(1, 43) = 0.05$, $p = .82$, $\eta^2 = .001$, nor the negative context, $F(1, 43) = 0.01$, $p = .97$, $\eta^2 < .001$.

To test whether the obtained context effects were uniquely related to Bob, automatic evaluations of unknown individuals were submitted to the same 3 (Context Valence) $\times 2$ (Time) ANOVA. This analysis revealed a significant main effect of time, $F(1, 43) = 4.74$, $p = .04$, $\eta^2 = .099$, showing that automatic evaluations of unknown individuals generally became less positive over time (*M*s = 0.57 vs. 0.50, respectively). Importantly, the crucial two-way interaction of Context and Time failed to reach statistical significance, $F(2, 86) = 1.76$, $p = .18$, $\eta^2 = .039$.

Discussion

The main goal of Experiment 3 was to test how heterogeneity of evaluative experiences influences automatic evaluations in novel, unfamiliar contexts. Our results suggested that automatic evaluations in novel contexts reflect the valence of the information acquired first. However, there was no evidence for the assumptions that automatic evaluations reflect the average or sum of all available information, or the valence of the information acquired last. Thus, it seems that earlier acquired information generalises to novel contexts even when subsequent information in a different context is evaluatively inconsistent with the initial information.

EXPERIMENT 4

Even though the results of Experiment 3 suggested that automatic evaluations in novel, unfamiliar contexts reflect the valence of the information acquired first, one could object that this pattern may emerge only when the initial information is negative, but not when it is positive. Research on negativity bias has repeatedly shown that negative information has a stronger impact than positive information (e.g., Cacioppo & Berntson, 1994; Fazio, Eiser, & Shook, 2004; Skowronski & Carlston, 1989). Thus, it seems possible that the results obtained in Experiment 3 were driven by a stronger impact of negative compared to positive information, rather than by a stronger impact of earlier compared to later acquired information. To address this question, Experiment 4 tested the influence of heterogeneity of evaluative experiences on automatic evaluations in novel contexts when initially acquired positive information was qualified by subsequently presented negative information. If the results obtained in Experiment 3 were driven by a stronger impact of negative information, automatic evaluations in novel contexts should change over the two learning sessions, with positive evaluations at time 1 and negative evaluations at time 2. If, however, the results obtained in Experiment 3 were driven by a stronger impact of initially acquired information, automatic evaluations in novel contexts should remain stable over the two learning sessions and in line with the positive evaluations predicted for the initial context.

Method

Participants and design. Forty-four undergraduates at the University of California, Santa Barbara, participated for research credit. The experiment consisted of a 2 (Order of Background: yellow first, blue second vs. blue first, yellow second) $\times$ 3 (Context During Measurement: yellow vs. white vs. blue) $\times$ 2 (Time of Measurement: time 1 vs. time 2) mixed-model design, with the first factor varying between-subjects and that latter two varying within-subjects.

Procedure. The evaluative learning paradigm was identical to the one employed in Experiment 3, the only difference being that participants in Experiment 4 were presented with positive information in the first learning session and negative information in the second learning session. As a measure of automatic evaluations, we again used Payne et al.'s (2005) affect misattribution procedure (AMP). The AMP was largely identical to the one employed in Experiment 3, the only difference being that we increased the overall number of trials from 90 to 150. Thus, each of the two face primes (i.e., Bob, unknown) was presented 25 times on each of the three backgrounds (i.e., yellow, blue, white). As with Experiments 1 and 3, automatic evaluations were assessed twice using the same measure, once after the first session of the evaluative learning task (time 1) and once after the second session (time 2).

Results

To examine context effects on automatic evaluations, we again calculated the proportion of *more pleasant* responses for each of the three types of primes (i.e., Bob–yellow; Bob–blue; Bob-white) at time 1 and time 2, respectively. Data were then aggregated as a function of context valence in the two learning sessions (i.e., positive vs. neutral vs. negative), collapsing across the method-factor of background colour.

Submitted to a 3 (Context Valence: positive vs. neutral vs. negative) × 2 (Time of Measurement: time 1 vs. time 2) mixed-model ANOVA, automatic evaluations of Bob revealed a significant main effect of Context Valence, $F(2, 86) = 4.44$, $p = .02$, $\eta^2 = .094$, and, more important, a significant two-way interaction of Context Valence and Time of Measurement, $F(2, 86) = 3.69$, $p = .03$, $\eta^2 = .079$ (see Figure 5). Replicating the pattern obtained in Experiments 1 and 3, automatic evaluations did not differ as a function of

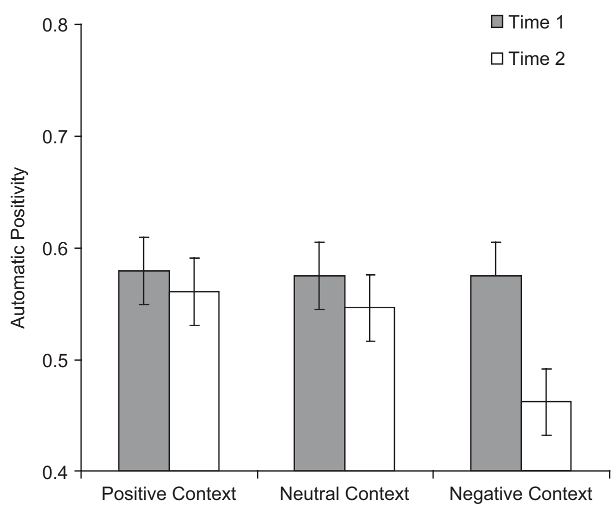

Figure 5. Automatic positivity toward the target individual as a function of context valence (positive vs. neutral vs. negative) and time (time 1 vs. time 2) when positive information was presented at time 1 and negative information at time 2, Experiment 4.

the context at time 1, $F(2, 86) = 0.03$, $p = .97$, $\eta^2 = .001$. In contrast, automatic evaluations significantly differed as a function of the context at time 2, $F(2, 86) = 5.25$, $p = .007$, $\eta^2 = .109$. Specifically, automatic evaluations were more positive in the positive context than in the negative context, $F(1, 43) = 8.91$, $p = .005$, $\eta^2 = .172$. More important for the present question, automatic evaluations in the neutral context significantly differed to automatic evaluations in the negative context, $F(1, 43) = 9.14$, $p = .004$, $\eta^2 = .175$, but not to automatic evaluations in the positive context, $F(1, 43) = 0.15$, $p = .70$, $\eta^2 = .004$. Moreover, automatic evaluations became more negative as a function of time in the negative context, $F(1, 43) = 8.69$, $p = .005$, $\eta^2 = .168$, but not in the neutral context, $F(1, 43) = 0.37$, $p = .55$, $\eta^2 = .009$, or positive context, $F(1, 43) = 0.21$, $p = .65$, $\eta^2 = .005$.

To test whether the obtained context effects were uniquely related to the target individual Bob, automatic evaluations of unknown individuals were submitted to the same 3 (Context Valence) × 2 (Time of Measurement) ANOVA. This analysis revealed a marginally significant main effect of Time, $F(1, 43) = 3.44$, $p = .07$, $\eta^2 = .072$, indicating that automatic evaluations of unknown individuals generally became more positive over time (Ms = 0.49 vs. 0.55, respectively). Importantly, the crucial two-way interaction of Context and Time failed to reach statistical significance, $F(2, 86) = 1.45$, $p = .24$, $\eta^2 = .033$.

Discussion

The main goal of Experiment 4 was to test whether the results obtained in Experiment 3 were due to (a) a stronger impact of negative as compared to positive information or (b) a stronger impact of earlier compared to later acquired information. For this purpose, Experiment 4 reversed the order of information in the evaluative learning paradigm, such that initially positive information was qualified by subsequently presented negative information. The results of Experiment 4 showed that automatic evaluations in novel contexts were driven by a stronger impact of initially acquired information. That is, automatic evaluations in novel contexts reflected the valence of earlier acquired information irrespective of whether this information was positive or negative. These findings are consistent with the claim that new information that contradicts a prior expectation may be stored as an "exception to the rule", such that the initial expectation is still regarded as an accurate default case. As such, evaluative information that is inconsistent with prior expectancies seems to determine subsequent automatic evaluations only in the context in which this information was acquired, with earlier acquired attitudes still generalising to novel contexts.

GENERAL DISCUSSION

Research has shown that automatic evaluations can be highly context-sensitive (see Blair, 2002; Gawronski & Bodenhausen, 2006, for reviews). Such contextual influences have been demonstrated for a wide range of different contexts, such as social settings (e.g., Wittenbrink et al., 2001), social roles (e.g., Barden et al., 2004), accessible exemplars (e.g., Dasgupta & Greenwald, 2001), salient categories (e.g., Mitchell et al., 2003), goal pursuit (e.g., Ferguson & Bargh, 2004), and food deprivation (e.g., Seibt et al., 2007). The main goal of the present research was to investigate basic principles guiding the formation of context-dependent and context-independent automatic attitudes. Specifically, we were interested in (a) the causal role of evaluatively heterogeneous experiences for context-dependent and context-independent automatic evaluations and (b) the principles associated with the contextualisation versus generalisation of automatic evaluations in novel contexts. Based on recent theorising (e.g., Gawronski & Bodenhausen, 2007), we argued that automatic evaluations of an object should be context independent (i.e., generalise to other contexts) when the associative representation of this object is evaluatively homogeneous. In contrast, automatic evaluations should be sensitive to contextual cues when the associative representation is evaluatively heterogeneous. In the latter case, automatic evaluations should reflect the contingency between the valence of previously learned information and the context in which this information was acquired. Consistent with these predictions, we found evidence for context-independent automatic evaluations that reflected the valence of previously acquired information when this information was evaluatively homogeneous; however, when the acquired information was evaluatively heterogeneous, automatic evaluations became context dependent, such that they reflected the valence of the information that was associated with a particular context during learning. Moreover, when information was evaluatively heterogeneous across contexts, novel contexts elicited automatic evaluations reflecting the valence of earlier acquired information, implying that earlier acquired experiences still generalise to novel contexts when experiences with attitude object are evaluatively heterogeneous.

Implications for attitude change

The present results have important implications for research on attitude change. Specifically, our findings provide further support for the notion that new experiences may not erase old associations from memory, but simply add new associations to the already existing representation (Gregg et al., 2006; Petty et al., 2006; Rydell & McConnell, 2006). In other words, old attitudes do not really go away, but simply become more complex when new evaluative information is acquired. Moreover, the present research shows that evaluatively heterogeneous representations, based on different experiences across different contexts, may be responsible for the high context sensitivity of automatic evaluations in previous research (see Blair, 2002; Gawronski & Bodenhausen, 2006, for reviews). Specifically, our findings suggest that automatic evaluations should be consistent across contexts when earlier experiences are evaluatively homogeneous. However, automatic evaluations should be context sensitive when experiences with an attitude object are evaluatively heterogeneous. Thus, the current work implies that modifications of automatic attitudes would be most pronounced when the context during the learning of new information is similar to the context during the acquisition of the original attitude.[4]

Another important finding is that automatic evaluations in novel contexts reflected the valence of earlier acquired information, rather than the sum or average of all available information or the valence of the information acquired last. This finding resembles earlier theorising, stating that old attitudes may influence automatic evaluations despite significant changes in deliberate evaluative judgements (e.g., Petty et al., 2006; Wilson et al., 2000). However, the present findings go beyond these theories in several respects. For instance, Wilson et al.'s (2000) dual-attitudes model claims that newly acquired attitudes do not overwrite old attitudes in memory, but can co-exist with old "implicit" attitudes. According to their dual-attitudes theory, old attitudes get activated automatically, whereas newly acquired attitudes require a high amount of cognitive effort to be retrieved from memory. The present findings qualify Wilson et al.'s (2000) theorising by showing that both old and newly acquired information can get activated automatically. Importantly, whichever information gets activated depends on the particular context, such that automatic evaluations reflect the valence of the information that was associated with a particular context during learning (Bouton, 2005). Nevertheless, automatic evaluations in novel contexts seem to reflect earlier experiences, which resembles research on context-dependent extinction in the learning literature (e.g., Bouton, 2004) and Wilson et al.'s (2000) claim that earlier experiences show some kind of "superiority" over later experiences.

Another implication of the present research is that automatic evaluations may be difficult to change entirely, given that evaluatively inconsistent information may often be contextualised. In many real-life contexts,

[4] Note, however, that the acquisition of evaluatively inconsistent information in the same learning context may not necessarily change the original evaluative response in novel contexts. Instead, it seems possible that novel contexts may still reflect the valence of the earlier acquired response, even when that response is changed for the learning context (see Bouton & Ricker, 1994).

experiences that may contradict an earlier acquired attitude occur in contexts that are different to the ones in which the original attitude has been formed (e.g., positive experiences with a Black person in a particular context contradicting a negative attitude toward African Americans that has been formed in a different context). As such, these experiences will likely be tied to the context in which they occurred. Nevertheless, it seems possible that new experiences that are made in the same context in which the original attitude has been formed may qualify automatic evaluations independent of the context. For instance, Petty et al. (2006) have shown that automatic evaluations in a given context reflect the average of early and later acquired information, when heterogeneous information was encoded in the same context. To the degree that these findings can be interpreted as a change of the original attitude, and given that novel contexts tend to reflect a generalisation of the original attitude, it seems possible that automatic evaluations in novel contexts will also reflect the new attitude. However, qualifying these speculations, an alternative possibility is that the acquisition of evaluatively inconsistent information in the same learning context produces changes only in that particular context, with novel contexts still reflecting the valence of the earlier acquired response. For instance, Bouton and Ricker (1994) found that conditioned stimuli still elicited the original conditioned response in novel contexts, even when the conditioned response has been eliminated by means of extinction in the original learning context. Future research may help to further clarify the principles of generalisation versus contextualisation when contradictory information is learned in the same context.

Open questions and future directions

The present results also imply several new questions for future research. One question concerns the role of clustered versus dispersed contexts for the generalisation of early experiences to novel contexts. As shown in Experiments 3 and 4, novel contexts tend to elicit automatic evaluations that reflect the valence of early experiences, in that subsequent experiences are contextualised and early experiences still generalise to novel contexts. However, to the degree that context-specific, highly dispersed experiences of heterogeneous valence do not imply a univalent "early" experience (e.g., Experiment 2), the nature of automatic evaluations in novel contexts remains an open question for evaluative learning in dispersed contexts. Similar to Petty et al.'s (2006) findings showing an averaging effect in a given context when inconsistent information was encoded in the same context, it seems possible that dispersed experiences may result in an averaging effect for novel, unfamiliar contexts (but see Bouton & Ricker, 1994). Future research adding a novel, neutral context to a setup similar to the one employed in Experiment 2 may help to answer this question.

Another open question concerns the role of perceivers' awareness of the contingency between the valence of a given experience and the presence of context cues in which the experience is made. Admittedly, the background colour manipulation employed in the present studies was fairly blatant, making it quite likely that participants consciously noticed the contingency between valence and background colour. Thus, it seems an interesting question whether contingency awareness is a necessary precondition for the obtained contextualisation to occur (see Hardwick & Lipp, 2000), or whether automatic attitudes can become context sensitive even when perceivers are not consciously aware of the contingency between a given context cue and the valence of the experience that is made with the attitude object in the presence of that cue. On the one hand, it seems possible that the obtained contextualisation of automatic evaluations occurs only when perceivers consciously consider a given context cue as relevant (i.e., predictive) for the valence of their evaluative experience. In this case, contextualisation should not occur for any random cue that just happens to co-occur with a given evaluative experience, but only for those cues that are consciously regarded as relevant. On the other hand, it is also possible that the obtained contextualisation of automatic evaluations is driven by lower-level associative processes. In this case, conscious awareness of the contingency between valence and a context cue may not be required for the obtained contextualisation to occur. Instead, any context cue that just happens to co-occur with an evaluative experience may be sufficient to create context-dependent automatic evaluations. Future research investigating the role of contingency awareness may help to clarify whether the contextualisation of automatic evaluations is driven by higher-order propositional or lower-level associative learning (see De Houwer, in press; Gawronski & Bodenhausen, 2006).

The present research also has some interesting parallels to the notion of occasion setting in the conditioning literature (see Schmajuk & Holland, 1998, for an overview). In general terms, occasion setting refers to "the potential of a stimulus to clarify the predictive value of an ambiguous cue" (Miller & Oberling, 1998, p. 3). Applied to the present research, the target person Bob may be regarded as an ambiguous cue, such that it "predicts" both positive and negative valence. Moreover, the background colours can be regarded as occasion setters in that the colour of the background clarifies whether the stimulus Bob "predicts" positive or negative valence. Even though the role of occasion setting has been extensively studied for signal learning (see Schmajuk & Holland, 1998), there are very few studies that have applied the notion of occasion setting to evaluative learning (Baeyens, Crombez, De Houwer, & Eelen, 1996; Baeyens, Hendrickx, Crombez, &

Hermans, 1998; Hardwick & Lipp, 2000). Interestingly, the evidence for occasion setting in these experiments is somewhat mixed, with some studies showing occasion setting for affective responses (Hardwick & Lipp, 2000) and others showing evaluative generalisation despite the presence of an occasion setter (Baeyens et al., 1996, 1998). Based on the available evidence, Hardwick and Lipp (2000)) argued that occasion setting in evaluative learning may be more likely to occur when (a) evaluative responses are assessed with indirect measures, (b) when contingency awareness is high, and (c) when the occasion setter and the evaluative information are presented in different modalities. Given that the paradigm employed in our studies met at least the former two requirements, our results are consistent with the possibility of occasion setting in evaluative learning.

Drawing on these considerations, we believe that an application of basic findings from the occasion-setting literature may provide deeper insights into processes of attitude formation, in particular the conditions of contextualisation versus generalisation. One example that we are currently exploring in our lab is the impact of counter conditioning. Research in the conditioning literature suggests that occasion setters (OS) are not directly associated with the US, as they do not provide information about the general presence of the US. Instead, OS are assumed to be associated with the relation between the CS and US. That is, OS indicate which type of US will follow the presence of the CS. This difference is also reflected in the present findings, showing that the effects of the background colours were limited to the target individual Bob, but did not generalise to other unknown individuals. However, in the present studies our conclusions were based on the interpretation of a null effect, which make them admittedly weak. In addition, the control faces in our measure of automatic evaluation differed from the target individual Bob in terms of their familiarity, which further undermines the conclusions that can be drawn from these data. A more stringent test of the occasion-setting function of context cues would be to subsequently manipulate the valence of the context cues in an evaluative conditioning paradigm (see De Houwer, Thomas, & Baeyens, 2001). According to the occasion-setting account, such a manipulation should influence automatic evaluations of the contexts alone, while leaving their capability to modulate responses to the CS unaffected. In other words, if a blue context signals that Bob is good and a yellow context signals that Bob is bad, then the two background colours should still elicit corresponding automatic evaluations of Bob, even when the blue background by itself is subsequently learned to be negative and the yellow background is learned to be positive. Such findings would provide more stringent evidence for the assumption that context cues indeed modulate automatic evaluations of the attitude object, as implied by the occasion-setting account.

An important conceptual question concerns the precise interpretation of *context*. In the present research, we used the term *context* to refer to the momentary presence of a contextual cue, in this case the particular background colour of the computer screen. Based on this conceptualisation, we argued that the contingency between a context cue and valence during learning influences the subsequent activation of evaluative associations in the presence of this cue. However, contexts in real-life environments are obviously much richer and much more complex, in that they comprise a large number of cues that may or may not be relevant to a person's momentary goals. Moreover, functionally similar contexts in real-life environments are rarely identical, as it was the case in our study. Instead, real-life contexts are at best similar to each other, which highlights the importance of encoding-related categorisation processes. For instance, two classrooms may be regarded as similar in that both can be categorised as classrooms. However, they could also be regarded as dissimilar on the basis of their furniture, location, or wall-paint. Thus, it remains an interesting question how momentary goals and encoding-related categorisation processes may influence the impact of contextual cues on evaluative learning and subsequent automatic evaluations.

Conclusion

In sum, the present studies provided important insights into the principles that guide the formation of context-dependent versus context-independent automatic evaluations. Specifically, our research showed that automatic evaluations were consistent across contexts when experiences with the attitude object were evaluatively homogeneous. In this case, different sets of object-related input stimuli seem to activate the same pattern of associations in memory, leading to generalised automatic evaluations across different contexts. If, however, experiences with the attitude object were evaluatively heterogeneous, automatic evaluations of that object became context sensitive. In this case, which pattern of associations gets activated seems to depend on the context in which the object is encountered and the contingency between this context and particular subsets of information in the learning of evaluative information. Moreover, when information was heterogeneous across contexts, novel contexts elicited automatic evaluations that reflected the valence of earlier acquired information, implying that earlier acquired experiences still generalise to novel contexts. In other words, automatic evaluations can be context dependent or context independent, which is determined by the relative heterogeneity of the underlying evaluative representation. In addition, the present findings suggest that new experiences with an attitude object do not erase old associations from memory. Instead, it seems that new experiences simply add new associations, making the evaluative representation of the attitude object more complex. Needless to say, these findings have important implications not only for theories of attitude change, but for any attempt to modify automatic evaluations in applied areas.

Manuscript received 10 March 2008
Revised manuscript received 1 July 2008
Manuscript accepted 17 July 2008
First published online 15 September 2008

REFERENCES

Anderson, N. H. (1971). Integration theory and attitude change. *Psychological Review*, *78*, 171–206.

Baeyens, F., Crombez, G., De Houwer, J., & Eelen, P. (1996). No evidence for modulation of evaluative flavour–flavour associations in humans. *Learning and Motivation*, *27*, 200–241.

Baeyens, F., Hendrickx, H., Crombez, G., & Hermans, D. (1998). Neither extended sequential nor simultaneous feature positive training results in modulation of evaluative flavor–flavor conditioning in humans. *Appetite*, *31*, 185–204.

Barden, J., Maddux, W. W., Petty, R. E., & Brewer, M. B. (2004). Contextual moderation of racial bias: The impact of social roles on controlled and automatically activated attitudes. *Journal of Personality and Social Psychology*, *87*, 5–22.

Bassili, J. N., & Brown, R. D. (2005). Implicit and explicit attitudes: Research, challenges, and theory. In D. Albarracín, B. T. Johnson, & M. P. Zanna (Eds.), *Handbook of attitudes and attitude change* (pp. 543–574). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Betsch, T., Kaufmann, M., Lindow, F., Plessner, H., & Hoffmann, K. (2006). Different principles of information integration in implicit and explicit attitude formation. *European Journal of Social Psychology*, *36*, 887–905.

Blair, I. V. (2002). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review*, *6*, 242–261.

Bouton, M. E. (2004). Context and behavioral processes in extinction. *Learning & Memory*, *11*, 485–494.

Bouton, M. E. (2005). Behavior systems and the contextual control of anxiety, fear, and panic. In L. F. Barrett, P. M. Niedenthal, & P. Winkielmen (Eds.), *Emotion and consciousness* (pp. 205–227). New York: Guilford Press.

Bouton, M. E., & Bolles, R. C. (1979). Contextual control of the extinction of conditioned fear. *Learning & Motivation*, *10*, 445–466.

Bouton, M. E., & King, D. A. (1983). Contextual control of the extinction of conditioned fear tests of the associative value of the context. *Journal of Experimental Psychology: Animal Behavioral Processes*, *9*, 248–265.

Bouton, M. E., & Ricker, S. T. (1994). Renewal of an extinguished responding in a second context. *Animal Learning & Behavior*, *22*, 317–324.

Bouton, M. E., & Schwartzentruber, D. (1986). Analysis of the associative and occasion setting properties of contexts participating in Pavlovian discrimination. *Journal of Experimental Psychology: Animal Behavioral Processes*, *12*, 333–350.

Bouton, M. E., & Schwartzentruber, D. (1989). Slow reacquisition following extinction: Context, encoding, and retrieval mechanisms. *Journal of Experimental Psychology: Animal Behavioral Processes*, *15*, 43–53.

Cacioppo, J. T., & Berntson, G. G. (1994). Relationship between attitudes and evaluative space: A critical review, with emphasis on the separability of positive and negative substrates. *Psychological Bulletin*, *115*, 401–423.

Conrey, F. R., & Smith, E. R. (2007). Attitude representation: Attitudes as patterns in a distributed, connectionist representational system. *Social Cognition*, *25*, 718–735.

Dasgupta, N., & Greenwald, A. G. (2001). On the malleability of automatic attitudes: Combating automatic prejudice with images of admired and disliked individuals. *Journal of Personality and Social Psychology*, *81*, 800–814.

De Houwer, J. (2003). The extrinsic affective Simon task. *Experimental Psychology*, *50*, 77–85.

De Houwer, J. (in press). The propositional approach to associative learning as an alternative for association formation models. *Learning & Behavior*.

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.

de Liver, Y., van der Pligt, J., & Wigboldus, D. (2007). Positive and negative associations underlying ambivalent attitudes. *Journal of Experimental Social Psychology*, *43*, 319–326.

Deutsch, R., & Gawronski, B. (2008). When the method makes a difference: Antagonistic effects on "automatic evaluations" as a function of task characteristics of the measure. *Manuscript submitted for publication*.

Devine, P. G. (2001). Implicit prejudice and stereotyping: How automatic are they? Introduction to the special section. *Journal of Personality and Social Psychology*, *81*, 757–759.

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

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

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

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

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, *50*, 229–238.

Ferguson, M. J., & Bargh, J. A. (2004). Liking is for doing: The effects of goal pursuit on automatic evaluation. *Journal of Personality and Social Psychology*, *87*, 557–572.

Ferguson, M. J., & Bargh, J. A. (2007). Beyond the attitude object: Implicit attitudes spring from object-centered contexts. In B. Wittenbrink & N. Schwarz (Eds.), *Implicit measures of attitudes: Procedures and controversies* (pp. 216–246). New York: Guilford Press.

Gawronski, B., & Bodenhausen, G. V. (2005). Accessibility effects on implicit social cognition: The role of knowledge activation versus retrieval experiences. *Journal of Personality and Social Psychology*, *89*, 672–685.

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. (2007). Unraveling the processes underlying evaluation: Attitudes from the perspective of the APE Model. *Social Cognition*, *25*, 687–717.

Gawronski, B., Cunningham, W. A., LeBel, E. P., & Deutsch, R. (2008). Affective priming as a measure of implicit preferences: Reliability depends on attention to relevant features in response interference tasks. *Manuscript submitted for publication*.

Gawronski, B., LeBel, E. P., & Peters, K. R. (2007). What do implicit measures tell us? Scrutinizing the validity of three common assumptions. *Perspectives on Psychological Science*, *2*, 181–193.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*, 4–27.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. K. L. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, *74*, 1464–1480.

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.

Hardwick, S. A., & Lipp, O. V. (2000). Modulation of affective learning: An occasion for evaluative conditioning? *Learning and Motivation*, *31*, 251–271.

Harris, J. A., Jones, M. L., Bailey, G. K., & Westbrook, R. F. (2000). Contextual control over conditioned responding in an extinction paradigm. *Journal of Experimental Psychology: Animal Behavioral Processes*, *26*, 174–185.

Johnston, L., & Hewstone, M. (1992). Cognitive models of stereotype change: III. Subtyping and perceived typicality of disconfirming group members. *Journal of Experimental Social Psychology*, *28*, 360–386.

Klauer, K. C., & Teige-Mocigemba, S. (2007). Controllability and resource dependence in automatic evaluation. *Journal of Experimental Social Psychology*, *43*, 648–655.

Kunda, Z., & Oleson, K. C. (1995). Maintaining stereotypes in the face of disconfirmation: Constructing grounds for subtyping deviants. *Journal of Personality and Social Psychology*, *68*, 565–579.

Kunda, Z., & Oleson, K. C. (1997). When exceptions prove the rule: How extremity of deviance determines the impact of deviant examples on stereotypes. *Journal of Personality and Social Psychology*, *72*, 965–979.

Miller, N., & Campbell, D. T. (1959). Recency and primacy in persuasion as a function of the timing of speeches and measurements. *Journal of Abnormal and Social Psychology*, *59*, 1–9.

Miller, R. R., & Oberling, P. (1998). Analogies between occasion setting and Pavlovian conditioning. In N. A. Schmajuk & P. C. Holland (Eds.), *Occasion setting: Associative learning and cognition in animals* (pp. 3–35). Washington, DC: American Psychological Association.

Mitchell, J. P., Nosek, B. A., & Banaji, M. R. (2003). Contextual variations in implicit evaluation. *Journal of Experimental Psychology: General*, *132*, 455–469.

Nosek, B. A., & Banaji, M. R. (2001). The go/no-go association task. *Social Cognition*, *19*, 625–666.

Payne, B. K., Cheng, S. 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., Govorun, O., & Arbuckle, N. L. (2008). Automatic attitudes and alcohol: Does implicit liking predict drinking? *Cognition and Emotion*, *22*, 238–271.

Payne, B. K., McClernon, J. F., & Dobbins, I. G. (2007). Automatic affective responses to smoking cues. *Experimental and Clinical Psychopharmacology*, *15*, 400–409.

Petty, R. E., Fazio, R. H., & Briñol, P. (Eds.). (in press). *Attitudes: Insights from the new implicit measures*. New York: Psychology Press.

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.

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

Rudman, L. A. (2004). Sources of implicit attitudes. *Current Directions in Psychological Science*, *13*, 79–82.

Rudman, L. A., Phelan, J. E., & Heppen, J. B. (2007). Developmental sources of implicit attitudes. *Personality and Social Psychology Bulletin*, *33*, 1700–1713.

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., Strain, L. M., Claypool, H. M., & Hugenberg, K. (2007). Implicit and explicit attitudes respond differently to increasing amounts of counter-attitudinal information. *European Journal of Social Psychology*, *37*, 867–878.

Schmajuk, N. A., & Holland, P. C. (1998). *Occasion setting: Associative learning and cognition in animals*. Washington, DC: American Psychological Association.

Schwarz, N. (2007). Attitude construction: Evaluation in context. *Social Cognition*, *25*, 638–655.

Seibt, B., Häfner, M., & Deutsch, R. (2007). Prepared to eat: How immediate affective and motivational responses to food cues are influenced by food deprivation. *European Journal of Social Psychology*, *37*, 359–379.

Skowronski, J. J., & Carlston, D. E. (1989). Negativity and extremity biases in impression formation: A review of explanations. *Psychological Bulletin*, *105*, 131–142.

Weber, R., & Crocker, J. (1983). Cognitive processes in the revision of stereotypic beliefs. *Journal of Personality and Social Psychology*, *45*, 961–977.

Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review*, *107*, 101–126.

Wittenbrink, B., Judd, C. M., & Park, B. (2001). Spontaneous prejudice in context: Variability in automatically activated attitudes. *Journal of Personality and Social Psychology*, *81*, 815–827.

Wittenbrink, B., & Schwarz, N. (Eds.). (2007). *Implicit measures of attitudes: Procedures and controversies*. New York: Guilford Press.