European Journal of Social Psychology
Eur. J. Soc. Psychol. **37**, 867–878 (2007)
Published online 18 September 2006 in Wiley InterScience
(www.interscience.wiley.com) **DOI**: 10.1002/ejsp.393

Implicit and explicit attitudes respond differently to increasing amounts of counterattitudinal information†

ROBERT J. RYDELL[1]*, ALLEN R. McCONNELL[2], LAURA M. STRAIN[2], HEATHER M. CLAYPOOL[2] AND KURT HUGENBERG[2]
[1]*Department of Psychology, University of California, Santa Barbara, USA*
[2]*Department of Psychology, Miami University, Oxford, USA*

Abstract

This research examined the processes by which explicit and implicit attitudes changed to systematically differing levels of counterattitudinal (CA) information. Explicit attitudes changed quickly in response to relatively small amounts of CA information, reflecting rule-based reasoning. On the other hand, implicit attitudes changed more slowly in the face of CA information, reflecting the progressive accretion of evaluation-attitude object pairings. Thus, explicit attitudes were extremely malleable and changed quickly when CA information was presented, however, implicit attitudes revealed a slow, linear change trajectory resulting from the on-going accrual of information about the attitude object. Implications for the processes underlying implicit and explicit attitudes are discussed.

Understanding how people respond to counterattitudinal (CA) information has a long tradition in social psychological research on attitude change (e.g., Petty & Wegener, 1998), prejudice reduction (e.g., Dovidio, Kawakami, & Gaernter, 2000), and person memory (e.g., Srull & Wyer, 1989). Although many established models account for how people respond to CA information (i.e., information inconsistent with the valence of the initial information about a target), most of this research is based on participants' explicit, controllable reports of their attitudes. However, recent attitudes work has made a distinction between implicit attitudes (i.e., attitudes for which people do not initially have conscious access and for which activation cannot be controlled) and explicit attitudes (i.e., attitudes that people can report and for which expression can be consciously controlled). Additionally, a great deal of

*Correspondence to: Dr Robert J. Rydell, Department of Psychology, University of California, Santa Barbara, CA 93106-9660, USA. E-mail: rydell@psych.ucsb.edu
†Portions of this work were submitted by Robert J. Rydell in partial fulfillment of the requirements for the degree Doctor of Philosophy in Psychology at Miami University.

Received 24 February 2006
Accepted 26 June 2006

research has shown that implicit and explicit attitudes predict different types of behaviors, with implicit attitudes predicting spontaneous, non-verbal behaviors and judgments (e.g., Hugenberg & Bodenhausen, 2003; McConnell & Leibold, 2001; Rydell & McConnell, in press), and explicit behaviors predicting controlled judgments and responses (e.g., Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997). Not only may implicit and explicit attitudes lead to and predict different outcomes (Wilson, Lindsey, & Schooler, 2000), but it has been suggested that they may also change due to different processes (Gawronski & Bodenhausen, 2006).

The current work expands on recent dual-process models of explicit and implicit attitudes, with a particular emphasis on how implicit and explicit attitudes can be changed by different processes in the face of CA information. Although the processes underlying explicit attitude change are relatively well understood (Eagly & Chaiken, 1993; Petty & Wegener, 1998), the processes underlying implicit attitude change are not (Devine, 2001; Gawronski & Bodenhausen, 2006; Gregg, Seibt, & Banaji, 2006). Indeed, researchers have just begun to focus attention on how implicit attitudes are formed, function, and change (Fazio & Olson, 2003).

Most research showing implicit attitude change has relied on modifying the social context or the target-relevant stimuli to temporarily change the accessibility of knowledge associated with a stigmatized group in memory (e.g., Barden, Maddux, Petty, & Brewer, 2004; Blair, 2002; Dasgupta & Greenwald, 2001; Gawronski & Bodenhausen, 2006). On the other hand, and more relevant to the present concerns, other research has shown that implicit attitudes seem quite resilient to change, unlike explicit attitudes (e.g., Gregg et al., 2006; Petty, Tormala, Briñol, & Jarvis, 2006). For example, Petty et al. showed that when people change their implicit and explicit attitudes (e.g., learning that a person conditioned to be positive is dissimilar to you) they associated uncertainty with the attitude object. This uncertainty was due to the inconsistency in valence between their initial attitudes and their changed attitudes, indicating that the initial attitude was still exerting impact on attitude-relevant processes. Also, Gregg et al. showed that people could easily form implicit and explicit attitudes toward two novel but competing groups that differed in valence. Although explicit attitudes did change in response to CA information (whether imagined or provided), implicit attitudes did not. This research showed that implicit attitudes, once formed, were relatively impervious to new information. The findings of Gregg et al. have added to the considerable confusion about how implicit attitudes change in the wake of new information. Because implicit evaluations seem to be malleable in some circumstances and basic tenants of learning state that evaluations should change when enough attitude inconsistent information is presented (McClelland, McNaughton, & O'Reilly, 1995; Olson & Fazio, 2001), how implicit attitudes respond to CA information is still unclear. Therefore, the current research addresses how implicit attitudes respond to systematically different amounts of CA information. In particular, this research explored attitude change involving a novel attitude object to examine the basic processes underlying implicit and explicit attitudes in the absence of any pre-existing or overarching expectancies about the target.

Our approach to tackling the question of how implicit attitude change is realized draws heavily on dual-process and memory models that account for how different types of information processing affect the integration of new information (McClelland et al., 1995; Sloman, 1996; Smith & DeCoster, 2000). Specifically, we differentiate between two systems of evaluation that are based on different modes of processing which allow for (a) the utilization of evaluations that are built up over time and (b) those that are responsive to small amounts of evaluation inconsistent information (McClelland et al., 1995). The first system of evaluation is the slow-learning system, which operates by gradually accruing attitude object-evaluation associations. This system is known as the slow-learning system because it takes a considerable number of paired associations (based on similarity and temporal contiguity) for knowledge to develop or change. If change occurs in this system, it occurs incrementally. In contrast,

DOI: 10.1002/ejsp

the second system of evaluation is the fast-learning system, which relies on rule-based logic and higher-order cognitive processes to operate. Because of its reliance on higher-order cognition, knowledge in the fast-learning system can develop and change rapidly through the employment of logic, rules, and syllogistic reasoning.

When applied to the attitudes literature, a systems of evaluation approach fits nicely with current social psychological conceptualizations of implicit and explicit attitudes (Rydell & McConnell, in press; Rydell, McConnell, Mackie, & Strain, in press; Wilson et al., 2000). The slow-learning system can potentially advance our understanding of how implicit attitudes form and operate and the fast-learning system fits with social psychologists' conceptualization of explicit attitudes as evaluations based on conscious thought that use cognitive resources in their production, revision, and expression (e.g., Fazio, 1995). Following this logic, several researchers have proposed that implicit attitudes and explicit attitudes are the products of different underlying processing modes that are distinct (e.g., Gawronski & Bodenhausen, 2006; Smith & DeCoster, 2000; Wilson et al., 2000). In the current research, we examined in detail if implicit attitude change could be accounted for by the slow-learning system and if explicit attitudes could be accounted for by the fast-learning system.

In support of a system of evaluation perspective, Rydell et al. (in press) recently demonstrated that explicit attitudes were formed in response to consciously available information whereas implicit attitudes formed in response to the valence of subliminally presented primes when both types of information were available. Consistent with a systems of evaluation account, the formation of implicit and explicit attitudes were independent of each other, with each reflecting the type of information (associative vs. verbal) assumed to influence the slow-learning and fast-learning systems, respectively. Although this experiment established a dissociation between the fast-learning and slow-learning systems of evaluation with respect to the type of information for which each system is differentially sensitive, it did not evaluate the fast versus slow nature of each system of evaluation.

More relevant to the current research, Rydell and McConnell (in press) had participants form either a positive or a negative attitude toward a target person using only verbal information. Then, some participants received no CA information, a small amount of verbal CA information (20 pieces), or a large amount of verbal CA information (100 pieces). In line with the fast-learning system, explicit attitudes toward the target individual changed quickly, even with very little CA information (e.g., 20 pieces). On the other hand, implicit attitudes only changed in response to a considerable amount of CA verbal information (e.g., 100 pieces). Thus, this work shows that implicit attitudes change in response to large amounts of CA information (whether associative or verbal), but not to small amounts of verbal information. Yet, some questions remain. For example, a systems of evaluation perspective would anticipate gradual, linear implicit attitude change in the face of CA information because implicit attitudes should be based on the totality of target-relevant information encountered. However, it is possible that implicit attitudes change in a non-linear fashion, and instead, suddenly change once a certain threshold is reached (Blair, 2002). Thus, it would be valuable to systematically vary the presentation of CA information to observe how implicit and explicit attitudes respond differently.

Our systems of evaluation perspective would anticipate differential change in implicit and explicit attitudes based on different amounts of CA information. If explicit attitude change is based on a fast-learning system of evaluation, then substantial explicit attitude change should occur when relatively little CA information is presented and quickly asymptote as the amount of CA information encountered increases. However, if the process through which implicit attitudes change relies on a slow-learning system of evaluation, a linear change in implicit attitudes should be observed as the amount of CA information encountered increases.

DOI: 10.1002/ejsp

METHOD

Participants

A sample of 132 undergraduates participated in return for research credit in their introductory psychology courses. Participants were randomly assigned to a receive one of six different levels of CA information: 0 CA trials (control; 0 CA), 20 CA trials (20 CA), 40 CA trials (40 CA), 60 CA trials (60 CA), 80 CA trials (80 CA), 100 CA trials (100 CA).

Procedure

Participants completed a learning task on the computer that was modeled after the procedure of Kerpelman and Himmelfarb (1971). Specifically, participants were told that they would be learning about a person named Bob. In the initial learning trials, participants read a sequence of 100 different behaviors performed by Bob. After each behavior, participants were asked to indicate whether the behavior was characteristic or uncharacteristic of Bob by pressing a computer key. Once participants responded, they were given feedback about whether the behavior was actually characteristic of Bob. The feedback during these initial trials portrayed Bob as uniformly positive where positive behaviors were characteristic of Bob (e.g., 'Bob continually yells at his wife in public') and negative behaviors were uncharacteristic of Bob (e.g., 'Bob fought against a discriminatory law that made renting difficult for minorities').

Next, participants were presented with 100 additional trials to introduce CA information (from the participants' perspective, these subsequent trials were simply a continuation of the first 100 trials). In these CA trials, participants continued to indicate whether each behavior presented was characteristic or uncharacteristic of Bob (as before). In the control condition (0 CA), there were 100 neutral trials (i.e., the behavior characteristic of Bob was neither positive nor negative in valence; e.g., 'Bob bought a shelf'). In the other five experimental conditions, the feedback provided about Bob was CA (i.e., negative behaviors were characteristics of him and positive behaviors were uncharacteristic of him) for either the next 20, 40, 60, 80, or 100 trials (for the 20 CA, 40 CA, 60 CA, 80 CA, or 100 CA conditions, respectively). Once all of the CA information was presented, any remaining trials (e.g., the last 80 trials in the 20 CA condition) provided neutral feedback (as in the 0 CA control condition). Thus, all participants received 200 pieces of information about Bob, with the first 100 trials indicating that Bob's behaviors were positive in nature, and the subsequent 100 trials presenting varying degrees of CA information. After presenting all 200 learning trials, participants completed measures of implicit and explicit attitudes.[1]

A subset of the positive and negative behaviors developed by McConnell, Sherman, and Hamilton (1994a) were used for both the initial learning and the subsequent CA learning trials (behaviors were only presented once and they were selected to ensure that behaviors in the initial 100 trials were not directly at odds with behaviors in the second 100 trials). A picture of Bob was presented on the computer monitor directly above each behavior during all learning trials.[2] Each behavior remained on the monitor until participants made a judgment as to whether the behavior was characteristic or

[1]Half of the participants completed the implicit measure first and the other half completed the explicit measure first. This order variable produced no effects in this experiment and is not discussed further.
[2]Photographs of five different White Males were randomly presented as 'Bob.' The photograph used did not affect the results in this experiment.

DOI: 10.1002/ejsp

uncharacteristic of Bob, and afterwards, they were informed (on the monitor for 5000 ms) that their judgment was either 'correct' or 'incorrect.'

Measure of Explicit Attitudes

After completing the learning task, participants judged how likable Bob was on a scale ranging from 1 (*very unlikable*) to 9 (*very likable*). In addition, they completed five semantic differential scales using 9-point scales to characterize Bob: good–bad, mean–pleasant, agreeable–disagreeable, uncaring–caring, and cruel–kind. Further, participants provided their attitude toward Bob on a feeling thermometer that ranged in temperature from 0 to 100°. The score for each explicit measure was standardized and then the scores were averaged to form the measure of explicit attitudes ($\alpha = 0.98$), with greater scores indicating more positive explicit attitudes toward Bob.

Measure of Implicit Attitudes

The Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998) was used to assess implicit attitudes toward Bob. The IAT task consisted of 26 stimuli: 1 picture of Bob, 5 different pictures of White Males who were not Bob, 10 positive adjectives (e.g., wonderful), and 10 negative adjectives (e.g., disgusting).

This IAT required participants to judge if a picture presented was Bob or Not Bob, or if a word presented was positive or negative in valence. On the critical trials, two category labels were mapped onto a single response key that were either attitude-consistent ('Bob or Positive' was paired on one key and 'Not Bob or Negative' was paired on a second key) or were attitude-inconsistent ('Bob or Negative' was paired on one response key and 'Not Bob or Positive' was paired on a second key). Words or pictures were randomly selected such that five positive and five negative words were presented once in each block, the picture of Bob was presented five times in each block, and each non-Bob picture was presented once per block.

Implicit attitudes were computed by subtracting the mean latency for the trials that paired 'Bob or Positive' and 'Not Bob or Negative' from the mean latency for the category trials that paired 'Bob or Negative' and 'Not Bob or Positive' and then dividing that difference score by the standard deviation of the latencies of all the trials in the two critical blocks, thus, creating Greenwald, Nosek, and Banaji's (2003) *D* measure. Thus, greater, positive scores on the IAT reflected more positive attitudes toward Bob. These *D* scores were then standardized with greater scores reflecting more positive attitudes toward Bob.

RESULTS

To examine if implicit and explicit attitudes differed in response to varying amounts of CA information, a 2 (standardized attitude scores: implicit attitude, explicit attitude) × 6 (CA condition: control, 20 CA, 40 CA, 60 CA, 80 CA, 100 CA) repeated measures ANOVA (with the former factor within-subjects) was conducted. Importantly, the results revealed the predicted two-way interaction, $F(5,126) = 2.22$, $p = 0.05$. To examine the interaction in detail, the simple effect of CA condition was examined for implicit and explicit attitudes separately and, when significant effects were observed, post-hoc comparisons were made using Tukey's Honestly Significant Difference.

DOI: 10.1002/ejsp

Explicit attitudes were examined using a one-way ANOVA of CA condition. The predicted main effect for CA condition obtained, $F(5,126) = 11.32$, $p < 0.001$, and showed that participants in the control condition had more positive attitudes toward Bob than participants in any other condition. Participants' explicit attitudes did not differ significantly among any of the other CA conditions (i.e., the 20 CA, 40, CA, 60 CA, 80 CA, 100 CA conditions did not differ), revealing the predicted asymptotic effect. Although explicit attitudes did show a significant linear trend, $F(1,126) = 28.48$, $p < 0.001$, the pattern of data across CA conditions did deviate significantly from linearity, $F(4,126) = 7.03$, $p < 0.001$. Additionally, both a quadratic and cubic trend significantly predicted explicit attitudes, $F(1,126) = 21.80$, $p < 0.001$ and $F(1,126) = 4.61$, $p < 0.04$, respectively. Thus, a linear trend does not best account for the explicit attitude data obtained and inspecting the pattern of data in Figure 1 makes this abundantly clear.

Implicit attitudes were also submitted to a one-way ANOVA of CA condition. The predicted main effect for CA condition obtained, $F(5,126) = 2.77$, $p < 0.03$, but as expected, the pattern was very different than it was for explicit attitudes. Specifically, implicit attitudes toward Bob became progressively more negative as a greater amount of CA information was presented (see Figure 2). Importantly, implicit attitudes showed a significant linear trend, $F(1,126) = 13.27$, $p < 0.001$, and this pattern did not deviate significantly from linearity, $F(4,126) = 0.14$, $p = 0.97$. Neither the quadratic nor cubic trends reached significance, $F < 1$. It is clear from these results and Figure 2 that implicit attitudes showed the predicted linear pattern of attitude change.

Although the explicit measures revealed an extremely high level of interrelation, we examined each of the explicit measures separately because four of the semantic differentials involved traits that may have been implied by some of our behaviors: mean–pleasant, agreeable–disagreeable, uncaring–caring, and cruel–kind. Thus, the results could be due to changes in the trait ratings and not in the measures that specifically target positive and negative evaluations (liking, good–bad, feeling thermometer). However, as can be seen from Table 1, regardless of the type of measure used in the explicit attitude composite, the same results obtained, namely a significant effect of CA condition, $F(5,126) > 3.89$, $p < 0.005$. In every case, the pattern of data matched the

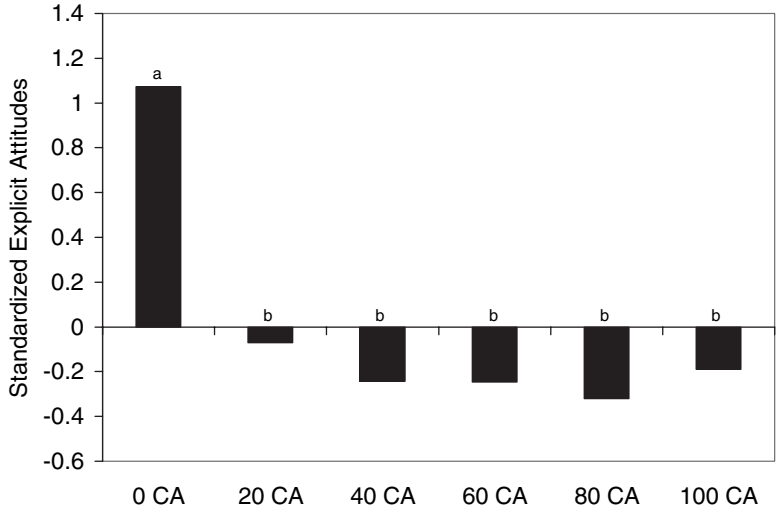

Figure 1. Explicit attitudes as a function of counterattitudinal condition. Note: Means (represented by the bars) of the figure that do not share a common alphanumeric label differ at the 0.05 level (Tukey HSD). CA, counterattitudinal information; 0 CA, control condition

DOI: 10.1002/ejsp

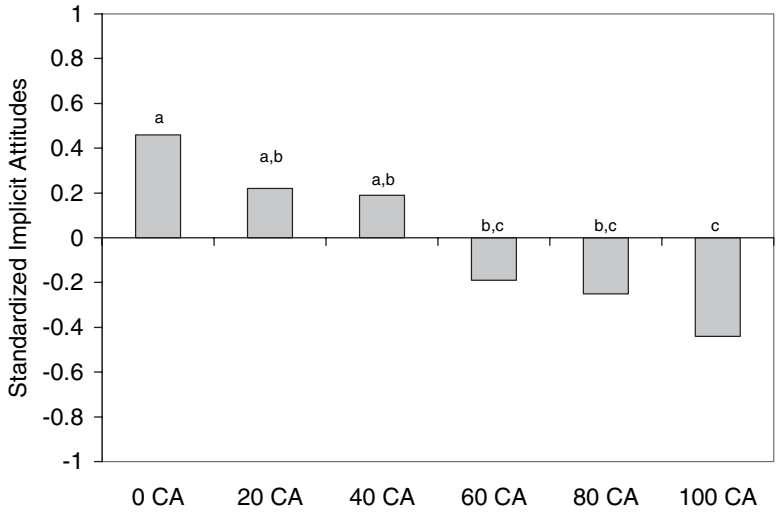

Figure 2. Implicit attitudes as a function of counterattitudinal condition. Note: Means (represented by the bars) of the figure that do not share a common alphanumeric label differ at the 0.05 level (Tukey HSD). CA, counterattitudinal information; 0 CA, control condition

pattern for the composite explicit measure, with only the 0 CA condition differing from all the other CA conditions. Therefore, it is extremely unlikely that explicit attitude results reported above were due to the traits implied by the behaviors, but instead were based on overall evaluations of Bob.

If implicit and explicit attitudes are the product of two different systems of evaluation, it may also be expected that there would be a weak or non-significant relation between implicit and explicit attitudes. In fact, implicit and explicit attitudes showed a marginally significant positive correlation, $r = 0.16$, $p < 0.08$. However, the magnitude of this correlation did not change as a function of CA condition (0 CA $r = 0.16$, ns, 20 CA $r = -0.06$, ns, 40 CA $r = -0.20$, ns, 60 CA $r = 0.25$, ns, 80 CA $r = 0.13$, ns, 100 CA $r = 0.05$, ns). This weak correlation tangentially supports conceptualizing implicit and explicit attitudes as changing via distinct mechanisms.

Table 1. Explicit attitude measures as a function of the amount of counterattitudinal (CA) information presented

	0 CA	20 CA	40 CA	60 CA	80 CA	100 CA
Liking	7.55_a	5.73_b	5.68_b	5.36_b	5.73_b	5.59_b
Semantic differentials						
Bad/good	8.05_a	5.90_b	5.54_b	5.68_b	5.54_b	5.59_b
Mean/pleasant	7.54_a	5.77_b	5.00_b	5.54_b	4.95_b	5.40_b
Disagreeable/agreeable	7.91_a	5.68_b	5.04_b	5.00_b	5.13_b	5.50_b
Uncaring/caring	8.18_a	5.82_b	5.77_b	5.63_b	5.23_b	5.77_b
Cruel/kind	7.90_a	5.59_b	5.18_b	5.55_b	5.27_b	5.54_b
Feeling thermometer	76.09_a	59.91_b	61.50_b	55.63_b	54.41_b	57.27_b

Note: Means in a row that do not share a common subscript differ at the 0.05 level (Tukey HSD). CA, counterattitudinal information; 0 CA, control condition.

DOI: 10.1002/ejsp

DISCUSSION

These findings support our contention that implicit and explicit attitudes can change consistent with different systems of evaluation. Explicit attitudes changed in response to small amounts of attitude-inconsistent information and showed an asymptotic trajectory, supporting our position that explicit attitudes are formed and changed by a fast-learning system of evaluation. However, implicit attitudes changed much more slowly in response to the same CA information. Importantly, implicit attitude change was linear and continued as more and more CA information was encountered, consistent with implicit attitudes being modified by a slow-learning, associative system of evaluation. Extending our past work, we showed detailed evidence of different processes involved in implicit and explicit attitude change by showing that they reveal different signatures of attitude change in response to increasing amounts of CA information. The current research provides evidence that strongly supports our supposition that implicit attitudes change slowly, in a linear fashion to increasing amounts of CA information.

Also consistent with a slow-learning system of evaluation and the basic tenants of several models of attitude change (Wilson et al., 2000), this work showed that implicit attitudes were changed by CA information when enough of it was encountered to offset the initial implicit attitude. Rather than conceiving of this outcome as the original implicit attitude 'being replaced' with a 'new attitude' as more CA information is encountered, we contend that the CA information provides additional associations that, as more and more are added, begin to change the overall implicit evaluation of the target. This research suggests that implicit and explicit attitudes were changed through different processes because the same information led to very different implicit and explicit evaluations of an attitude object (McClelland et al., 1995; Smith & DeCoster, 2000).[3]

This research goes beyond that of Rydell and McConnell (in press; see also Rydell et al., in press) because it elucidates, in greater detail, the process by which implicit attitudes slowly change in response to new information about an attitude object. Specifically, as more and more negative CA information was encountered, implicit attitudes became more and more negative. This extension is important because it tests an assumption of the system of evaluation framework that implicit attitude change is observed in response to large amounts of CA information because implicit attitudes reflect the slow accrual of target-relevant information (see also McClelland et al., 1995; Smith & DeCoster, 2000). Indeed, this is exactly what we observed in the current work.

Although this research speaks more directly to how implicit attitudes change in response to CA information, the results for explicit attitudes were also intriguing. Most importantly, the current work showed that explicit attitudes did not change significantly in response to greater amounts of CA information once a small amount of CA information was encountered (i.e., from the 20 CA condition to the 100 CA condition). In other research, we have observed the same pattern of explicit attitude change (see also, Kerpelman & Himmelfarb, 1971) but this past work did not vary the target-relevant feedback in a fashion needed to map out its signature across greater amounts of CA information. Moreover, Rydell and McConnell (in press) showed that when people receive initially clear-cut information about a target person (i.e., the target goes from always being positive to always being negative), they respond to the initial change in the target person's behavior, adjust their explicit attitudes accordingly and quickly, and then pay very little attention to additional CA information about the target's subsequent actions. However, when people receive a relatively mixed initial pattern of target behavior (e.g., the target is sometimes positive and sometimes negative but then later is always negative), perceivers were more likely to 'suspend immediate judgments of the target' rather than rush to form a strong on-line

[3]However, it should be noted that this does not preclude implicit ambivalence in response to an attitude object caused by the initial evaluation of that attitude object (Petty et al., 2006).

DOI: 10.1002/ejsp

impression (cf., McConnell, Sherman, & Hamilton, 1994b), and accordingly, attended to more subsequent CA information. This likely occurs because perceivers expect considerable consistency in the behavior of an individual (McConnell, 2001), and thus, continue to work effortfully to understand individuals who violate this expectation (i.e., those who initially reveal mixed behaviors). Thus, we would not always expect explicit attitudes to asymptote after such a small amount of information. Rydell and McConnell (in press) showed that explicit attitudes more closely track the amount of CA information when people are motivated to process CA information extensively (i.e., they receive initially mixed information toward an attitude object or are given a goal to attend to later information conscientiously) as suggested by traditional models of explicit attitude change (Petty & Wegener, 1998).

Further, this study only examined how positive attitudes can be changed by negative CA information, yet it is possible that the reverse order (i.e., learning positive information following initial negative information about a target) would produce weaker outcomes because negative information is perceived as more diagnostic and is more influential in forming impressions about liking (Skowronski & Carlston, 1987). Although we would not expect the signature of implicit and explicit attitudes to deviate from the current linear and asymptotic trajectories (respectively), we might expect the steepness of these patterns to be diminished when positive behaviors come on the heels of initially negative information about the target. Consistent with this reasoning, Rydell and McConnell (in press) found that negative CA information led to greater explicit attitude change than did positive CA information; however, these valence asymmetries were not revealed for implicit attitudes (see also, Carlston & Skowronski, 2005). Because extracting a trait from behavior may rely on some amount of effortful processing (Bassili & Smith, 1986) and may require verbal processes as well (Carlston, 1994), valence asymmetries seem to occur for explicit attitudes but not for implicit attitudes. Based on our past research, we would expect to obtain the same results for implicit attitudes if initial learning was negative, but we would have expected to see less dramatic explicit attitude change.

Also, given that the different implicit and explicit attitudes can simultaneously exist, one can imagine how different reactions toward attitude objects could be exhibited for different types of behavior. Research has shown that implicit attitudes are much more predictive of spontaneous, subtle behaviors whereas explicit attitudes are much more predictive of deliberate, intentional behaviors (e.g., Dovidio et al., 1997; McConnell & Leibold, 2001). Although behavioral data were not collected in the current study, we would predict interesting behavioral asymmetries would be exhibited toward the attitude object, especially under conditions where the greatest divergence in implicit and explicit attitudes is revealed. Specifically, people may exhibit inconsistent behaviors when those behaviors are controllable (and presumably guided by explicit attitudes) versus uncontrollable (and presumably guided by implicit attitudes) at the same time (e.g., McConnell & Leibold).

The current work is consistent with many tenets of attitude models that envision implicit and explicit attitudes as products of different underlying evaluative systems (e.g., Gawronski & Bodenhausen, 2006; Smith & DeCoster, 2000; Wilson et al., 2000). But, it should be noted that the systems of evaluation approach outlined in this manuscript is, at times, inconsistent with prominent models of attitudes (Eagly & Chaiken, 1993; Fazio, 1995). For example, the MODE model (Fazio) does not perceive differences between implicit and explicit attitudes as evidence of people holding different attitudes, but rather, as showing differences in the extent to which people have the motivation (e.g., motivation to control prejudice) and ability (e.g., cognitive resources) to modulate an attitude automatically activated from memory (Fazio & Olson, 2003). The current research does not speak directly to the proposed differences between these models, but instead, is primarily focused on exploring the basic processes whereby implicit and explicit attitudes change in response to CA

DOI: 10.1002/ejsp

information. As such, we should make it clear that the present research can only inform these different models, not differentiate among them.

In addition, the current work uses measures of implicit and explicit attitudes to study the processes involved in implicit and explicit attitude change. Research conducted in this manner confounds measurement with process in that the form of measurement (i.e., implicit or explicit) is used to elucidate process (i.e., automatic or controlled). Because the measures are different in several ways, this leads to some ambiguity in interpretation of process. Although there is, at present, no foolproof way to avoid this confound (especially in a single study), the results should be interpreted with this limitation in mind. It should be noted that although this limitation exists in many different research programs, we attempted to minimize this confound by using a novel attitude object for which people were not highly motivated to evaluate in a certain way and for whom future social interaction was not expected. These steps should reduce the likelihood that social desirability or self-presentational motives might cloud the validity of the current measures.

Also, because the current work was interested in exploring the basic processes underlying implicit and explicit attitude change, it may have somewhat limited generalizability. For instance, it is not often the case that in everyday situations a person would go from always acting positively to suddenly acting negatively. However, there are situations in which people can, in one domain of their life, act extremely positively, but in another domain act extremely negative. For example, consider a colleague who is great around the office but quite despicable at home. If one came to know this colleague more intimately and became introduced to her home life, how would you view her? One's overall explicit evaluation may be able to differentiate the colleague at work and at home (e.g., McConnell, Leibold, & Sherman, 1997), but on an implicit level, how would this new information affect evaluation? Based on the current research, it seems likely that implicit attitudes would become more negative in response to the individual only when learning much more information about her behavior at home. In addition, Briñol, Petty, and Wheeler (2006) have shown that implicit–explicit attitude discrepancies like those observed in the current study can increase processing of discrepancy-related information. Because people can often have discrepant implicit and explicit attitudes and these discrepancies have cognitive implications, understanding differences between implicit and explicit attitude change and potentially reducing or increasing these discrepancies would be important for understanding ourselves and others.

CONCLUSIONS

Elucidating differential processing behind implicit and explicit attitude change is extremely important for advancing theoretical conceptualizations of attitude change. Thus, this research provides useful insights about differences in implicit and explicit attitude change and could be used to examine several unaddressed issues in the attitude change literature more directly. Specifically, this knowledge may be important for understanding why attitudes, especially long-established attitudes built upon many associations in memory, are so difficult to change over long periods of time, or why at other times, people's opinions can change quickly.

ACKNOWLEDGEMENTS

This research was supported by NIMH grant MH068279 and NSF grant BCS 0516931.

DOI: 10.1002/ejsp

REFERENCES

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., & Smith, M. C. (1986). On the spontaneity of trait attribution: Converging evidence for the role of cognitive strategy. *Journal of Personality and Social Psychology*, *50*, 239–245.

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

Briñol, P., Petty, R. E., & Wheeler, S. C. (2006). Discrepancies between explicit and implicit self-concepts: Consequences for information processing. *Journal of Personality and Social Psychology, 91*, 154–170.

Carlston, D. E. (1994). Associated Systems Theory: A systematic approach to cognitive representations of persons. In R. S. Wyer, Jr (Ed.), *Advances in social cognition* (Vol. 7, pp. 1–78). Hillside, NJ: Erlbaum Associates.

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

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.

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.

Dovidio, J. F., Kawakami, K., & Gaernter, S. L. (2000). Reducing contemporary prejudice: Combating explicit and implicit bias at the individual and intergroup level. In S. Oskamp (Ed.), *Reducing prejudice and discrimination* (pp. 137–163). Mahwah, NJ: Erlbaum.

Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the nature of prejudice: Automatic and controlled processes. *Journal of Experimental Social Psychology*, *33*, 510–554.

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Orlando: Harcourt, Brace, and Jovanovich.

Fazio, R. H. (1995). Attitudes as object-evaluation associations: Determinants, consequences, and correlates of attitude accessibility. In R. E. Petty, & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 247–282). Mahwah, NJ: Erlbaum.

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

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.

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

Greenwald, A. G., Nosek, B. A., & Banaji, M. H. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*, 197–216.

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

Hugenberg, K., & Bodenhausen, G. V. (2003). Facing prejudice: Implicit prejudice and the perception of facial threat. *Psychological Science*, *14*, 640–643.

Kerpelman, J. A., & Himmelfarb, S. (1971). Partial reinforcement effects in attitude acquisition and counter-conditioning. *Journal of Personality and Social Psychology*, *19*, 301–305.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *102*, 419–457.

McConnell, A. R. (2001). Implicit theories: Consequences for social judgments of individuals. *Journal of Experimental Social Psychology*, *37*, 215–227.

McConnell, A. R., & Leibold, J. M. (2001). Relations among the Implicit Association Test, discriminatory behavior, and explicit measures of racial attitudes. *Journal of Experimental Social Psychology*, *37*, 435–442.

McConnell, A. R., Leibold, J. M., & Sherman, S. J. (1997). Within-target illusory correlations and the formation of context-dependent attitudes. *Journal of Personality and Social Psychology*, *73*, 675–686.

McConnell, A. R., Sherman, S. J., & Hamilton, D. L. (1994a). Illusory correlations in the perception of groups: An extension of the distinctiveness based account. *Journal of Personality and Social Psychology*, *67*, 414–429.

DOI: 10.1002/ejsp

McConnell, A. R., Sherman, S. J., & Hamilton, D. L. (1994b). On-line and memory-based aspects of individual and group target judgments. *Journal of Personality and Social Psychology*, *67*, 173–185.
Olson, M. A., & Fazio, R. H. (2001). Implicit attitude formation through classical conditioning. *Psychological Science*, *12*, 413–417.
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.
Petty, R. E., & Wegener, D. T. (1998). Attitude change: Multiple roles for persuasion variables. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., Vol. 1, pp. 323–390). New York: McGraw-Hill.
Rydell, R. J., & McConnell, A. R. (in press). Understanding implicit and explicit attitude change: A systems of reasoning analysis. *Journal of Personality and Social Psychology.*
Rydell, R. J., McConnell, A. R., Mackie, D. M., & Strain, L. M. (in press). Of two minds: Forming and changing valence inconsistent implicit and explicit attitudes. *Psychological Science*.
Skowronski, J. J., & Carlston, D. E. (1987). Social judgment and social memory: The role of cue diagnosticity in negativity, positivity, and extremity biases. *Journal of Personality and Social Psychology*, *52*, 689–699.
Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, *119*, 3–22.
Smith, E. R., & DeCoster, J. (2000). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review*, *4*, 108–131.
Srull, T. K., & Wyer, R. S. (1989). Person memory and judgment. *Psychological Review*, *96*, 58–83.
Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review*, *107*, 101–126.

DOI: 10.1002/ejsp