

2010-02

YOU NEVER GET A SECOND CHANCE TO MAKE A FIRST (IMPLICIT) IMPRESSION: THE ROLE OF ELABORATION IN THE FORMATION AND REVISION OF IMPLICIT IMPRESSIONS

Wyer, NA

<http://hdl.handle.net/10026.1/2947>

10.1521/soco.2010.28.1.1

SOC COGNITION

GUILFORD PUBLICATIONS INC

All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Please cite only the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.

YOU NEVER GET A SECOND CHANCE TO MAKE A FIRST (IMPLICIT) IMPRESSION: THE ROLE OF ELABORATION IN THE FORMATION AND REVISION OF IMPLICIT IMPRESSIONS

Natalie A. Wyer
University of Plymouth

Conditions under which implicit and explicit impressions of an individual may change in response to new information were investigated in two experiments. Participants formed an impression of a target person based on his membership in a social group and, in some conditions, detailed behavioral evidence. Later, half of the participants were given reason to believe that the initial information they had been given was wrong, and that the target actually belonged to a different social group. Implicit and explicit measures of participants' impressions of the target were then collected. Results indicated that, while explicit impressions were effectively corrected in light of new information, implicit impressions continued to reflect initial beliefs (Experiments 1 and 2). However, when given the opportunity to re-examine the original behavioral information, implicit measures also reflected a change in participants' impressions (Experiment 2). The role of elaboration in determining implicit and explicit impression change is discussed.

The notion that first impressions last forever is a veritable cliché. The scientific evidence, however, is mixed. Indeed, research seems to indicate that *explicit* impressions about others are actually fairly easy to change, as long as the basis for the original belief is undermined, discredited, or counteracted in some way (e.g., Golding, Fowler, Long, & Lotta, 1990; Wyer & Unverzagt, 1985). On the other hand, a growing body of evidence suggests that *implicit* impressions are less flexible. Research indicates that initial beliefs (trait impressions or evaluations) can often be detected using implicit measures, even when explicitly reported beliefs have changed (e.g., Gregg, Seibt, & Banaji, 2006; Petty, Tormala, Brinol, & Jarvis, 2006;

Au: Change to
Spelling of
"Latta" ok?
↑



a
^

The author would like to thank Sabine Pahl and several anonymous reviewers for their comments on an earlier version of this manuscript. This research was presented at the 8th annual meeting of the Society for Personality and Social Psychology in Memphis, TN.

Correspondence concerning this article should be addressed to Natalie A. Wyer, School of Psychology, University of Plymouth, Drake Circus, Plymouth PL4 8AA, UK; E-mail: natalie.wyer@plymouth.ac.uk.

Rydell, McConnell, Mackie, & Strain, 2006; Wilson, Lindsey, & Schooler, 2000). The present research extends prior work by examining conditions that foster the retention of implicit impressions based on a person's group membership once new information contradicting that membership has been learned. In particular, this article will focus on the role of elaboration in both the formation and the revision of stereotype-based impressions.

UNDER WHAT CONDITIONS DO WE REVISE EXPLICIT AND IMPLICIT IMPRESSIONS?

There are a number of situations that potentially instigate a change in one's beliefs about another person or group. In the literature on explicit and implicit attitudes and impressions, there has been little consistency in how such conditions are defined, thus it is worth considering evidence for belief change across the range of situations in which it may occur.¹

Do Impressions Change When New Information Is Learned? Perhaps the most common paradigm for investigating changes in how a person or group is perceived involves providing participants with initial information (on the basis of which they form an impression or evaluation) and then presenting new information that is inconsistent with it. Early research on impression formation (see Hamilton & Sherman, 1996 for a review) suggested that explicit impressions are resistant to change, as perceivers find ways of explaining away impression-inconsistent information. More recent research, however, suggests that explicit evaluations are actually easier to change than implicit ones. For example, Rydell et al. (2006) obtained evidence that explicit attitudes toward a target person were readily formed and changed in response to behavioral information. As perceivers encountered new evidence that contradicted earlier information, they adjusted their explicit evaluations accordingly. In contrast, implicit attitudes toward the same target person were impervious to behavioral evidence (driven instead by subliminally presented primes).

Further evidence is provided by McConnell, Rydell, Strain, and Mackie (2008) in their research on implicit evaluations of members of stigmatized groups. Explicit attitudes toward both stigmatized and non-stigmatized targets were responsive to the amount of positive or negative behavioral information that described them. Likewise, initially negative implicit evaluations of non-stigmatized individuals became more positive following large amounts of positive information (and more negative following negative information). In contrast, targets identified as members of a stigmatized group continued to evoke negative implicit evaluations, regardless of the positivity of behaviors attributed to them.

¹ Research on the malleability of implicit attitudes toward real groups (e.g., Blair, Ma, & Lenton, 2006; Dasgupta & Asgari, 2004) arguably provides another case of belief change. However, in that research, responses to the group are believed to be based on a temporarily salient subset of the group, and thus do not provide evidence on the question of whether or not beliefs about the group as a whole can be changed. For this reason, this line of research is excluded from the present discussion. Likewise, research on the issue of whether group stereotypes are implicitly applied to group members (e.g., Kunda, Davies, Adams, & Spencer, 2002) is beyond the scope of this discussion.

Au: Hamilton & Sherman, 1996 not in refs.

Au: Not in refs.

Au: Date change ok?

2001
A

Finally, even when a target does not belong to a stigmatized out-group, implicit attitudes appear to be relatively resistant to change. Rydell and McConnell (2006) reported that moderate amounts of counterattitudinal information were sufficient to reverse participants' explicit, but not implicit, evaluations of a target. Only under conditions in which perceivers received large quantities of counterattitudinal information did their implicit attitudes "catch up" with their explicit beliefs. Thus, evidence that is sufficient to update explicit beliefs appears to be inadequate to provoke perceivers to revise implicit evaluations (see also Rydell, McConnell, Strain, Claypool, & Hugenberg, 2007).

Au: Not in refs. →
Do you mean
Rydell, McConnell,
Mackie, & Strain,
2006?

Do Impressions Change When the Target Changes Over Time? A second situation in which beliefs might be expected to change is when a target's characteristics undergo an actual change. This possibility was investigated by Gregg et al. (2006, experiment 4), who found that although participants' explicit evaluations of two groups were easily revised in response to information that the groups' characters had changed, their implicit evaluations of the same groups continued to reflect the groups' original characteristics. Even when participants received compelling evidence that a previously disliked group had reformed and become quite likeable, participants' originally negative implicit attitudes persisted. Thus, like the situation described above, where additional information contradicts an initial impression, Gregg et al.'s (2006, Experiment 4) findings indicate that actual changes in a target's character do not translate into changes in how it is implicitly evaluated.

Au: Not in refs.

Do Impressions Change When Erroneous Information is Corrected? A final situation in which beliefs about a person or group might reasonably be expected to change—and the one under direct investigation in the present work—occurs when perceivers learn that information that they received earlier is actually false. Although early research on belief perseverance suggested that perceivers persist in believing discredited information (e.g., Ross, Lepper, & Hubbard, 1975), later research made it clear that explicit beliefs do change when previously learned information is unambiguously false (Golding et al., 1990; Wyer & Unverzagt, 1985).

Au: Not in refs. →

More recently, attention has turned to the question of whether *implicit* impressions also change when previously held beliefs are explicitly discredited. A study by Wilson, Lindsey, and Aronson (1998, cited in Wilson et al., 2000) found that participants continued to hold negative implicit attitudes toward a target who they initially believed to be a child molester but later learned was a lawyer prosecuting the child molester. Such findings contradict parallel studies on explicit belief perseverance in that implicit attitudes persisted even when the basis for them had been discredited. Further evidence in support of this possibility is provided by Gregg et al. (2006, Experiment 3) who told participants that information previously learned about two groups had been switched such that the group about which they had learned negative information was actually positive, and the group about which they had learned positive information was actually negative. Despite explicitly endorsing the "correct" evaluations of the two groups, participants' implicit evaluations reflected the information they had originally been given. Thus, even in situations where perceivers explicitly believe information to be false, that information may still influence their implicit impressions or evaluations.

WHY DO IMPLICIT IMPRESSIONS PERSIST?

A number of theories have emerged in recent years to account for differences between implicit and explicit evaluations. Some such theories focus on distinguishing between multiple memory systems (e.g., Devine, 1989; Evans & Over, 1996; Smith & DeCoster, 2000; Wilson et al., 2000) while others emphasize the role of distinctive types of information processing (e.g., Gawronski & Bodenhausen, 2006; Sloman, 1996; Strack & Deutsch, 2004). While there are important differences among these theories, they share the assumption that, under different conditions, responses may reflect either automatic associations with the target or more deliberately retrieved propositional knowledge about the target.

Of particular interest here is the proposition by several theories (e.g., Gawronski & Bodenhausen, 2006; Petty et al., 2006) that deliberative responses are more likely to reflect perceivers' understanding of the "truth" of a belief, whereas automatic responses are less influenced by a belief's validity. For example, according to Gawronski and Bodenhausen's associative-propositional evaluation (APE) model, responses based on propositional (relatively deliberative) processing are heavily influenced by the "truth value" (p. 693) of the information considered. In contrast, associative (relatively automatic) responses are generated "irrespective of whether a person considers these . . . to be accurate or inaccurate" (p. 693). In other words, associative processing produces automatic or intuitive reactions which may or may not coincide with what one believes about the stimulus, whereas propositional processes deliver responses that are necessarily held to be accurate.

Similarly, Petty et al.'s (2006) "past attitudes are still there" (PAST) model posits that a prior attitude is "tagged" as being false once it has been changed. However, in order for a new attitude to override the previous one, it must be stronger than the prior attitude and the prior attitude's "false" tag must be retrieved. In cases where the prior attitude is stronger, or where the tag is weak, the prior attitude is likely to carry more weight. In Petty et al.'s (2006, Study 1) research, explicit attitudes toward two targets reflected information that participants had recently learned about them. However, implicit attitudes were more neutral (or ambivalent, as described by Petty et al.) when that information contradicted previously classically conditioned evaluations. That is, on an implicit measure that did not allow for retrieval of a "false" tag, prior and current attitudes carried equal weight and cancelled each other out.

THE ROLE OF ELABORATION

Although developed to provide an account of attitude change, such models may apply equally well to other types of beliefs (e.g., person impressions). While dual attitudes resulted in relatively neutral/ambivalent implicit attitudes in Petty et al.'s (2006) study, the PAST model suggests that the nature of implicit responses should depend on three factors: (1) the strength of the initial attitude, (2) the strength of the new attitude, and (3) the strength of the association between the initial attitude and its "false" tag. Each of these components is likely to be influenced by the extent to which one has elaborated on the relevant information. For

example, Petty et al. (2006) note that participants in their research (Study 1) did not actively reject their initial attitudes, hence the association between those attitudes and their “false” tags was likely weak. Strengthening this association (by encouraging perceivers to actively reject their prior beliefs) might be expected to result in a decreased influence of prior beliefs on implicit responses.

Likewise, manipulations that alter the strength of either prior beliefs or current beliefs should also influence implicit responses, as it is primarily the relative strength of these beliefs that determines which will be expressed. The strength of a belief is influenced by the extent to which one elaborates upon it (Petty & Cacioppo, 1986). Thus prior beliefs that have been subject to more extensive elaboration should be more likely to affect implicit responses than those that have not been elaborated upon. Likewise, current beliefs are more likely to influence implicit responses to the extent that one has extensively elaborated upon them.

The notion that the strength of one’s beliefs is determined by the amount of elaboration that has contributed to them recalls early research by Srull and Wyer (1979, 1980) who found that activating a trait concept had little direct effect on judgments of a target person—it was only when the activated trait could be used to interpret ambiguous information about the target that it biased impressions in trait-consistent ways. Similar findings were reported by Kunda and Sherman-Williams (1993) who found that accessible stereotypes determined judgments of a target only through their biasing effect on how ambiguous information was construed (see also Banaji, Hardin, & Rothman, 1993). Both of these lines of research are consistent with the idea that accessible traits or stereotypes have their greatest influence on impressions by shaping the way subsequently learned information is elaborated upon.

AW: Srull & Wyer, 1979 and 1980 not in refs.

AW: These 2 refs not listed.

EXPERIMENT OVERVIEW AND HYPOTHESES

The current experiments aimed to investigate the role of elaboration in the formation and revision of stereotype-based beliefs about others. A secondary aim was to expand on prior research on implicit beliefs which has focused on *evaluations* of individuals (Petty et al., 2006; Rydell et al., 2006) or groups (Gregg et al., 2006). In the current research, implicit *stereotypic* beliefs about an individual (rather than simple evaluations per se) will be the focus. In the two studies presented here, participants formed an impression of an individual based on photographs and written information. The information was provided in two stages. First, participants learned that the individual belonged to a particular social category (i.e., skinhead or cancer patient). Some participants also received ambiguous behavioral evidence that could be interpreted in terms of either category and thus were able to elaborate on the category information provided. Next, all participants learned that the individual actually belonged to the same category (i.e., cancer patient). Thus, participants who had initially learned that the individual was a skinhead had reason to revise their impressions, whereas those who initially learned that he was a cancer patient had no reason to do so. In Experiment 2, some participants were asked to review the behavioral evidence they had previously learned in light of the new category information, and thus were given an opportunity to elaborate on

that information. After learning about the individual, participants completed both explicit and implicit measures of their impressions of him.

EXPERIMENT 1

Experiment 1 tested the hypothesis that prior beliefs would only influence implicit impressions when those beliefs were elaborated upon. Because explicit responses should be readily updated, it was expected that they should reflect current beliefs regardless of either prior beliefs or behavioral evidence. Previous work by Gregg et al. (2006, Experiment 3) suggests that explicit evaluations of two groups were readily changed by participants who learned that the initial information they had been given was incorrect. However, the same participants failed to revise their implicit attitudes toward the groups, suggesting that implicit evaluative associations remained a part of their group representations. This finding is consistent with the reasoning outlined above. Participants were able to elaborate on their initial evaluations in the course of learning extensive information about the groups. In contrast, the fact that the original information was incorrect received relatively little elaboration.

Experiment 1 builds on the results reported by Gregg et al. (2006) in that it investigates explicit and implicit impressions of an individual (rather than group) target. In addition, stereotype-based impressions will be assessed rather than simple evaluations. Most importantly, however, Experiment 1 includes a direct manipulation of elaboration in order to test the hypothesis that differences in the amount of elaborative processing will result in differences in the extent to which implicit impressions are resistant to change.

METHOD

PARTICIPANTS

Eighty female undergraduate students at the University of Plymouth (average age 22.2 years) took part in this and an unrelated experiment in order to fulfill a course requirement. Participants were tested individually or in non-interacting groups of two to five.

DESIGN

The experiment employed a 2 X 2 design, with *category change* (category change or no change) and *behavioral evidence* (evidence or no evidence) manipulated between-participants.

MATERIALS AND PROCEDURE

Experimental instructions and stimuli were presented via computer using E Prime software. Participants learned that they would read about a man named Edward first by viewing photographs of him and then by reading about recent events in which he was involved. Participants then viewed six color photographs of the same man in various contexts (at home, at a restaurant, at a party, etc.). The man appeared to be approximately 25 years of age and had a clearly visible shaved head. In all photographs, the man was casually dressed and showed a neutral facial expression. Participants were asked to familiarize themselves with Edward's appearance by looking closely at each photograph.

Participants were then told that they would learn more about Edward by reading information about him. They were asked to imagine that they were introduced to Edward at a party by a mutual acquaintance. The information that followed varied as a function of *category change* condition, in order to manipulate the initial category to which Edward was assigned. Participants in the *no change* condition were told that the acquaintance informed them (after Edward had walked away) that Edward was suffering from a form of cancer and was undergoing chemotherapy, which had led to the loss of his hair. Participants in the *change* condition were told that the acquaintance informed them that Edward was a skinhead and had shaved his head after joining a local gang.

Participants were also assigned to one of two behavioral evidence conditions. Those in the *evidence* condition were told to imagine that, over the course of the party, they observed Edward on a number of occasions and noted a number of things about his behavior. The description was constructed to be ambiguous, but generally consistent with the stereotypes of both skinheads and cancer patients. For example, participants learned that Edward was thin and appeared tired and haggard, which could be viewed as evidence that he parties a lot or that he is ill. They further learned that he was inattentive when interacting with others, which could be taken a sign of rudeness or of his not feeling well. Participants in the *no evidence* condition were not given the description to read.

Next, all participants were told to imagine that, the following day, they met another friend for coffee and mentioned their encounter with Edward. The friend informed them that they knew Edward well, and that he was in fact undergoing treatment for Hodgkin's disease, which had led to a dramatic change in his appearance. For participants in the *change* condition, this constituted a change to the information previously received.

DEPENDENT MEASURES

Participants completed both explicit and implicit measures of their impressions of Edward. First, participants were asked to rate Edward on 16 personality traits, including 8 positive and 8 negative traits, half of which were, according to pilot tests, related to the skinhead stereotype (*aggressive, unfriendly, hostile, rude, passive, social, kind, polite*, the last four being reverse-scored) and half of which were not

(*foolish, lazy, clumsy, boring, intelligent, hard-working, athletic, generous*). Traits were presented in a random order and participants were asked to indicate the extent to which each trait described Edward using a 5-point scale (1 = not at all to 5 = extremely).

After completing the trait ratings, participants were given instructions for an Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998), which provided an implicit measure of their impressions of Edward.²

In the IAT, participants classified a series of word and picture stimuli according to one of two dimensions, which varied across five blocks of trials, by pressing one of two response keys (Q and P). The words included six hostile words (*aggressive, violent, hostile, cold, intimidating, offensive*) and six kind words (*peaceful, kind, warm, approachable, friendly, polite*). The pictures included six pictures of Edward and six pictures of different men who were not Edward, but who were similar in appearance.

In Block 1, participants classified 24 photographs as Not Edward (Q) or Edward (P). In Block 2, participants classified 24 words as Kind (Q) or Hostile (P). In Block 3, participants classified 48 pictures and words as Kind/Not Edward (Q) or Hostile/Edward (P). In Block 4, participants classified 24 words as Hostile (Q) or Kind (P). In Block 5, participants classified 48 pictures and words as Hostile/Not Edward (Q) or Kind/Edward (P). The order of Blocks 2/3 and Blocks 4/5 was counterbalanced across participants.³ Within each block, the order of stimulus presentation was random.

The rationale underlying the IAT is that, to the extent that two concepts (e.g., Edward and hostile) are associated, one should be faster when using the same hand to respond to them. Thus, if Edward was associated with hostility (central to the stereotype of skinheads), participants should be faster to complete Block 3 than they are to complete Block 5.

After completing the IAT, participants were asked whether they believed that Edward was a cancer patient and/or a skinhead. Finally, after completing an unrelated experiment, they were debriefed and paid for their participation.

RESULTS AND DISCUSSION

PRETESTING OF TARGET DESCRIPTION

Prior to main experiment, the description of Edward was pretested to ensure that participants would (explicitly) view him as hostile when they believed him to be a skinhead but not when they believed him to be a cancer patient. Thirty-two

2. The IAT has been widely used to assess automatic associations of various kinds, including attitudes (Greenwald et al., 1998), self-concept (Greenwald & Farnham, 2000; Nosek, Banaji, & Greenwald, 2002), and other group-attribute associations (e.g., Dasgupta & Asgari, 2004). It is used here in favor of other implicit measures (e.g., the Go/No-Go Association Task, Nosek & Banaji, 2001) in order to allow for direct comparisons with previous research on implicit belief change (Gregg et al., 2006; Rydell et al., 2006).

3. Block order did not affect the results and will not be discussed further.

students at the University of Plymouth (20 females, average age 24.9 years) were shown photographs of Edward and were given either the "skinhead" or "cancer patient" description of him as outlined above. Half of the participants in each group were also given behavioral evidence as described above. Participants then immediately completed the explicit trait-rating measure.

Average ratings on hostility-related traits were subjected to a two-way (category X behavioral evidence) Analysis of Variance (ANOVA), which yielded only a significant main effect of category, $F(1, 31) = 24.18, p < .001, \eta_p^2 = .46$. Participants who believed Edward was a skinhead rated him as more hostile $M = 3.67, s = 0.31$ than did those who believed him to be a cancer patient $M = 2.95, s = 0.52$. This was true both when they had received behavioral evidence, $M = 3.88, s = .22$ vs. $M = 2.98, s = 0.50, t(14) = 4.63, p < .001$, and when they had not, $M = 3.47, s = 0.24$ vs. $M = 2.92, s = .57, t(14) = 2.49, p < .03$. Finally, behavioral evidence led to higher hostility ratings for those who believed Edward to be a skinhead, $t(14) = 3.53, p < .01$, but not for those who believed him to be a cancer patient, $t(14) = .23, ns$.

On the basis of these data, the category labels "skinhead" and "cancer patient" appear to yield the expected differences in perceived hostility. Providing behavioral evidence in addition to the category label appears to provoke greater perceptions of hostility when the category encourages interpretation of that behavior as hostile (as in the "skinhead" condition), but not when the category provides an alternative interpretation (as in the "cancer patient" condition).

MANIPULATION CHECK

Participants' beliefs about Edward's category membership were inspected to ensure that they explicitly categorized him as a cancer patient. Participants' responses indicated that 100% believed him to be a cancer patient and 0% believed him to be a skinhead.

MEASURING EXPLICIT AND IMPLICIT IMPRESSIONS

Participants' ratings of Edward on hostility-related traits were significantly intercorrelated ($\alpha = .76$) thus they were averaged to form a single explicit impression core. Implicit impression (IAT) scores were computed following the procedure used by Gregg et al. (2006). Response times (RTs) less than 300 ms or greater than 3000 ms were deemed outliers and removed from the analysis (approximately 2.9% of critical trials). Average RTs for Blocks 3 and 5 were then computed on the basis of all critical trials, including those corresponding to errors (approximately 9.1% of critical trials). Implicit impression scores were then computed by subtracting the average RT on Block 3 from the average RT on Block 5.⁴

Effects of category change and behavioral evidence on implicit and explicit impressions following McConnell et al. (2008), both explicit and implicit impression scores were standardized before being entered into a repeated measures ANOVA with

B

AU! Not in refs.

type:
B head

4. Alternative methods for computing IAT scores (e.g., Greenwald, Nosek, & Banaji, 2003; Rydell et al., 2006) produced equivalent results. Gregg et al.'s (2006) method produced the highest internal consistency (split-half $r_s = .76$ and $.67$ for Blocks 3 and 5 respectively) of the three methods.

category change and behavioral evidence included as between-participants factors. This analysis produced a significant main effect of category change, $F(1, 76) = 5.65, p < .03, \eta_p^2 = .07$, as well as significant two-way interactions between category change and measure, $F(1, 76) = 4.22, p < .04, \eta_p^2 = .06$, and between category change and behavioral evidence, $F(1, 76) = 4.07, p < .05, \eta_p^2 = .05$. However, as predicted, these effects were qualified by a significant three-way interaction, $F(1, 76) = 5.08, p < .03, \eta_p^2 = .06$ (see Figure 1).

To decompose the three-way interaction, separate category change X behavioral evidence ANOVAs were carried out on the implicit and explicit impression measures. Analysis of implicit impression ratings revealed a significant main effect of category change, $F(1, 79) = 11.77, p < .01, \eta_p^2 = .13$, which was qualified by a significant interaction between category change and behavioral evidence, $F(1, 79) = 10.52, p < .01, \eta_p^2 = .12$. Paired comparisons confirmed that, among participants who received behavioral evidence, those who initially believed Edward to be a skinhead produced stronger associations between Edward and hostility than did those who initially believed him to be a cancer patient, $t(38) = 4.47, p < .001$. In contrast, among participants who received no behavioral evidence, association between Edward and hostility were unaffected by the group to which he was initially attributed, $t(38) = .14, ns$. Analysis of participants' explicit impressions revealed no significant effects (all $F_s < 1$).

Finally, implicit and explicit measures were not significantly correlated, $r = .02, ns$. Within-cell correlations ranged from $-.38$ to $+.18$, smallest $p = .10$.

DISCUSSION

Consistent with prior research (e.g., Gregg et al., 2006), Experiment 1 provides evidence that implicit and explicit impressions of others differ in the extent to which they are amenable to change. However, unlike past studies, this experiment demonstrated that the extent to which implicit impressions are resistant to change is moderated by whether or not perceivers elaborate on the initial information they receive (in this case, information regarding the target's group membership). When impressions were based on a simple category label, both implicit and explicit measures reflected participants' updated beliefs about the target. In contrast, when category information was supported by behavioral evidence, participants' implicit and explicit impressions diverged under conditions in which prior and current beliefs conflicted. In that case, implicit impressions were more strongly influenced by prior beliefs whereas explicit impressions reflected current (corrected) beliefs. This finding suggests that relatively superficial impressions may be more susceptible to change than those based on more extensive processing.

EXPERIMENT 2

It is important to note that elaboration in Experiment 1 was induced by providing more extensive behavioral information about the target, which could be interpreted in line with a salient category. There are, of course, other ways in which to elaborate upon category information. However, in the present context, the elabo-

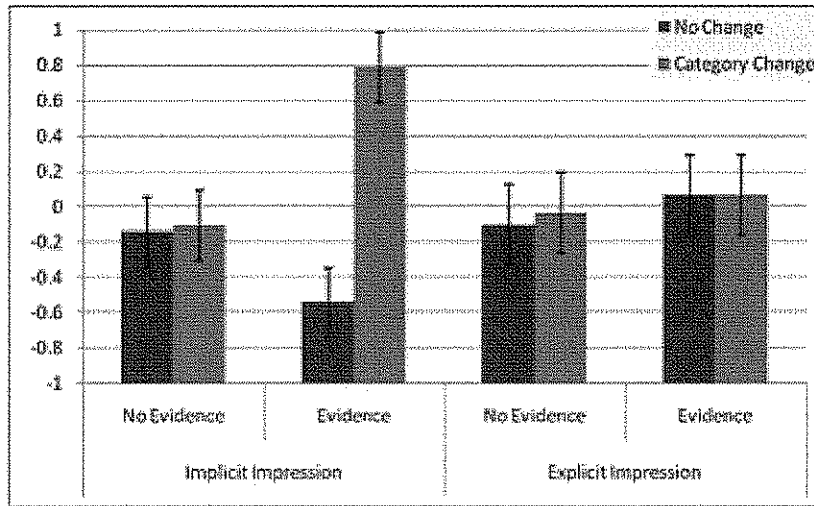


FIGURE 1. Implicit and Explicit impressions (Experiment 1) as a function of category change and behavioral evidence, based on standardized IAT scores and standardized trait ratings.

ration that took place—that is, biased processing of ambiguous behavioral information—may have led to impressions of the target that were in fact relatively independent of the initial category information. In other words, participants formed implicit impressions that Edward was hostile not because he was a skinhead but because the category label “skinhead” altered the way in which subsequent information was interpreted. Consequently, replacing the skinhead category label with one unassociated with hostility had little effect on implicit impressions, because the biased interpretation of Edward’s behavior was retained and was not directly challenged (e.g., Trope & Alfieri, 1997).

AU: Not in refs. →

This raises the question of how perceivers’ original interpretations of a target’s behavior might be challenged and thus promote changes in implicit impressions. One possibility is that elaboration plays an additional role in the revision of person impressions. As described earlier, Petty et al.’s (2006) PAST model suggests that when a person changes his or her attitude toward a target, the prior attitude coincides with the new one, but is tagged as “false.” Which attitude is expressed depends in part on whether the “false” tag is retrieved along with the prior attitude at the time of response. Retrieval of the “false” tag depends, in turn, on the strength of its association with the prior attitude. In the context of revising an earlier impression, the strength of the “false” tag associated with an initial impression may depend on the extent to which one has elaborated on new information. For example, one may elaborate on a new category label by reinterpreting what one already knows about a target in light of that new information. This proposal is akin to that made by Strack and Mussweiler (2001), who argued that correction processes in social judgment often involve “recomputation”—that is, in generating a new judgment about a target that one has previously judged, one may start over from scratch, reconsidering all currently available evidence and forming a judgment anew on that basis.

AU: Not in refs. →

Thus, Experiment 2 was designed to extend Experiment 1 by investigating the role of elaboration in the revision of implicit impressions. It was expected that updated information about a target's category membership would override prior beliefs when participants were given the opportunity to reinterpret previous information in light of the new category. As in the first experiment, it was expected that explicit impressions should reflect current beliefs regardless of whether perceivers had initially held different beliefs or whether they had engaged in elaboration at any stage of impression formation or revision.

METHOD

PARTICIPANTS

Participants were 102 members of the Plymouth community (66 females,⁵ average age 27.6 years) who took part in this and an unrelated experiment in exchange for £3 (approximately \$6). Participants were tested individually or in noninteracting groups of two to six.

DESIGN

The experiment employed a 2 X 3 design in which *category change* (change or no change) and *behavioral evidence* (no evidence, evidence, or review evidence) were manipulated between-participants.

MATERIALS AND PROCEDURE

The materials and procedures for participants in the *evidence* and *no evidence* conditions were identical to those described in Experiment 1. Participants in the *review evidence* condition went through the same procedures as those in the *evidence* condition; however, after finally learning that Edward was a cancer patient, these participants were asked to re-read the behavioral information they had previously received about Edward before completing the dependent measures.

RESULTS AND DISCUSSION

MANIPULATION CHECK

Participants' beliefs about Edward's category membership were inspected to ensure that they explicitly categorized him as a cancer patient. Participants' responses indicated that 100% believed him to be a cancer patient and 0% believed him to be a skinhead.

5. Participant gender did not affect any results and thus will not be discussed further.

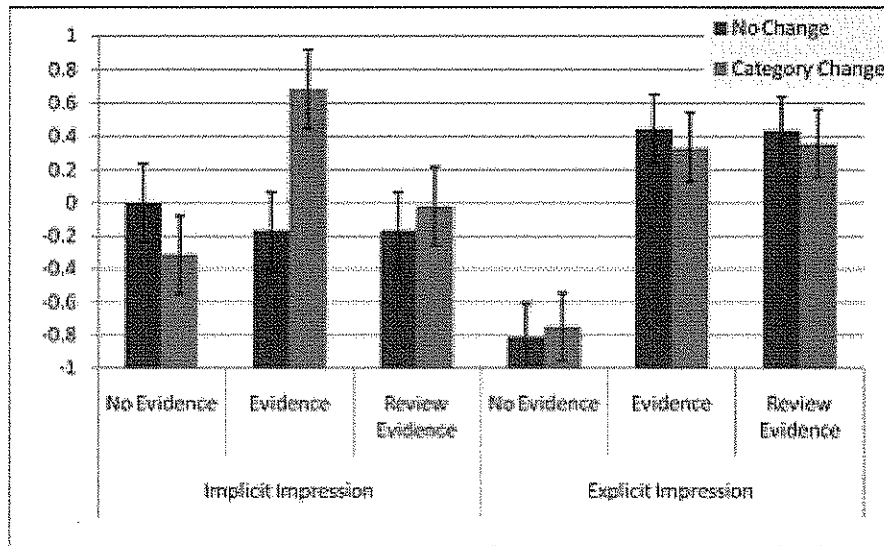


FIGURE 2. Implicit and Explicit impressions (Experiment 2) as a function of category change and behavioral evidence, based on standardized IAT scores and standardized trait ratings.

MEASURES OF EXPLICIT AND IMPLICIT IMPRESSIONS

Participants' ratings of Edward on hostility-related traits were significantly inter-correlated ($\alpha = .74$) thus they were averaged to form a single explicit impression score. Implicit impression scores were calculated as in Experiment 1.

EFFECTS OF CATEGORY CHANGE AND BEHAVIORAL EVIDENCE ON EXPLICIT AND IMPLICIT IMPRESSIONS

As in Experiment 1, explicit and implicit impression scores were standardized before being entered into a category change X behavioral evidence X impression measure mixed-model ANOVA, with repeated measures on the last factor. This analysis produced a significant main effect of behavioral evidence, $F(2, 96) = 13.827, p < .001, \eta_p^2 = .22$, as well as a significant evidence X measure interaction, $F(2, 96) = 5.49, p < .01, \eta_p^2 = .12$. Consistent with the hypotheses, these effects were moderated by a marginal three-way interaction, $F(2, 96) = 2.33, p = .10, \eta_p^2 = .05$ (see Figure 2).

In order to test the predicted effects on implicit and explicit impressions, separate category change X behavioral evidence ANOVAs were carried out. Analysis of implicit impressions yielded only a significant two-way interaction, $F(2, 96) = 3.09, p = .05, \eta_p^2 = .06$. Paired comparisons within each behavioral evidence condition confirmed that initial information about Edward's group membership influenced only participants who had received behavioral evidence but who had not had the opportunity to re-evaluate it, $t(32) = 2.43, p < .03$, leading those who initially learned that Edward was a skinhead to associate him with hostility to a greater extent than those who initially learned he was a cancer patient. Changes in Edward's group membership had no effect on participants who had received no behavioral evidence, $t(32) = 1.18, p > .20$. Furthermore, participants who had

received behavioral information but who had been given the chance to reassess it were unaffected by their initial beliefs about Edward's group membership, $t(32) = .41, p > .60$. Analysis of participants' explicit impressions yielded only a significant main effect of behavioral evidence, $F(2, 96) = 21.48, p < .001, \eta_p^2 = .31$. Bonferroni post-hoc tests indicated that participants in the *evidence* ($M = .39, s = .85$) and *review evidence* ($M = .39, s = .79$) rated Edward as more hostile than did those in the *no evidence* condition ($M = -.78, s = .87$), both $ps < .001$.

Finally, implicit and explicit measures were found to be uncorrelated with each other, $r = +.07, p = .50$. Within-cell correlations ranged from $-.19$ to $+.18$, smallest $p > .40$.

DISCUSSION

Experiment 2 extended the first study in several important ways. First, it provides further support for the role of elaboration in moderating the susceptibility of implicit (but not explicit) impressions to change. As in Experiment 1, implicit and explicit impressions responded in different ways to new information that contradicted them. Whereas explicit impressions consistently reflected participants' current beliefs about the target's category membership, implicit impressions continued to reflect their original beliefs—but only if those beliefs were supported by behavioral evidence. Second, Experiment 2 demonstrated that re-elaborating on information previously interpreted in terms of an original belief resulted in changes to implicit impressions. In contrast to those who merely learned that Edward belonged to a different category, participants who were allowed to revisit the same behavioral evidence they had received earlier no longer showed evidence of associating Edward with hostility. These results suggest that elaboration on a category label—in the form of using that label to interpret ambiguous information—plays two important roles in the formation and revision of implicit impressions. First, implicit impressions based on such elaboration are relatively resistant to change (whereas those based on relatively superficial processing are easily revised). Second, once implicit impressions are formed, opportunities for re-elaboration of the same information on which those impressions are based play a central role in changing them. Thus, this research bridges work on the revision of implicit attitudes or impressions with that on correction processes in explicit judgments (e.g., Strack & Mussweiler, 2001).

Figure 2 illustrates that implicit impressions of Edward were similar for participants in the no evidence and review evidence conditions. In both conditions, participants' associations between Edward and hostility were similar, regardless of whether they had initially believed him to be a skinhead—that is, the category change manipulation appears to have had no effect on implicit impressions. However, consideration of the evidence condition provides insight into how the no evidence and review evidence conditions may actually differ. The information to which participants were exposed in each of the three conditions was systematically varied such that participants in the evidence condition had access to the same information as was presented in the no evidence condition, and additional information (i.e., the behavioral evidence). Likewise, participants in the review evidence condition had access to the same information as was presented in the evidence condition, with the addition that they could re-read the behavioral information.

Att: Not in refs. →

In this sense, the three conditions represent three "snapshots" of the impression formation and revision process. It is therefore reasonable to assume that, had participants in the review evidence condition been assessed ^{ed} before they reread the behavioral information, their responses would have mimicked those of participants in the evidence condition. Thus, while the implicit impressions of participants in the no evidence and review evidence conditions appear, on the surface, to be similar, they have likely come about through very different routes.

One unexpected outcome of Experiment 2 was the finding that exposure to behavioral evidence led to higher explicit ratings of hostility, regardless of participants' beliefs about Edward's category membership and the extent to which they had elaborated on those beliefs. This is in contrast to Experiment 1 where behavioral evidence had no significant effect on explicit judgments of Edward, as well as the pretest to Experiment 1, where behavioral evidence only increased perceived hostility when Edward was believed to be a skinhead (but not when he was explicitly believed to be a cancer patient, as was the case in all conditions of Experiment 2). One possible explanation, albeit post hoc, is that the two samples differed in their level of motivation pertaining to the experiment. Whereas participants in Experiment 1 were undergraduate psychology students who were accustomed to participating in experiments and did so as part of a course requirement, those in Experiment 2 were members of the community at large who were largely inexperienced in the role of experimental participant and took part in exchange for monetary compensation. The latter group may have been more likely than the former to respond to the rating task according to their perceptions of what was expected (i.e., they were given unflattering information about Edward, and thus may have perceived that they were expected to evaluate him negatively). Responses to the implicit measure were, of course, less influenced by participants' level of compliance motivation. Although other plausible explanations may exist, the fact that the same pattern was not seen in Experiment 1 suggests that the main effect of behavioral evidence found in Experiment 2 was possibly anomalous or, at most, due to demand characteristics to which only highly motivated participants were susceptible.

GENERAL DISCUSSION

As the experiments reported here demonstrate, the manner in which perceivers initially categorize others can have a lasting influence on their implicit impressions, even when explicit beliefs about category membership change. Consistent with previous research (Gregg et al., 2006; Petty et al., 2006; Rydell et al., 2006), participants were successful in updating their explicitly-reported impressions of a target whose category membership had changed. However, when impressions were assessed using implicit measures, initial beliefs about the target's category membership continued to influence impressions under some conditions. Thus, like other types of beliefs (e.g., attitudes) stereotype-based impressions may continue to influence judgments even after the category information on which those impressions are based has been disregarded.

The results of these studies highlight the role of elaboration in determining the relative influence of prior and current beliefs. Participants failed to update their implicit associations between the target and the initial category to which he was

→ Au: Change to "assessed" ok?
↑

ed
^

assigned, but only when they had received behavioral evidence that reinforced those associations. Thus elaboration led to stronger initial beliefs, and therefore a stronger influence of initial beliefs on implicit responses.

In addition to its impact on the strength of initial beliefs, elaboration also appeared to influence the likelihood that revised beliefs could override the influence of their predecessors. Participants who were given the opportunity to elaborate on the new category information they received (by reviewing the behavioral evidence a second time) appeared to be successful in eliminating the influence of prior beliefs on implicit responses. Thus, the present research is generally supportive of the PAST model (Petty et al., 2006) in that participants were able to override their initial beliefs when either (1) those initial beliefs were weak or (2) those initial beliefs were actively rejected and replaced with well-elaborated (i.e., strong) new beliefs.

It is worth noting that elaboration in the present context takes the form of interpreting ambiguous behavioral information in light of a salient group stereotype (e.g., Banaji et al., 1993; Kunda & Sherman-Williams, 1993; Srull & Wyer, 1979, 1980). Elaboration may, of course, take other forms as well. For example, the attitude change literature conceives of elaboration in terms of self-generated responses to information (Petty & Cacioppo, 1986). Similar self-generated elaborations may also influence the impression formation (and revision) process. Social perceivers may, at times, spontaneously elaborate on category-based information (as when one actively infers traits, preferences, or abilities on the basis of a group stereotype). In such cases, implicit impressions may become more entrenched, compared to cases in which one merely notes a target's group membership without actively drawing inferences.

IMPLICIT BELIEFS—WHEN AND HOW ARE THEY CHANGED?

The present research adds to a growing body of work that demonstrates that implicit beliefs are more malleable than originally suspected (see Blair, 2002). A number of studies have now established that implicit out-group prejudices can be modified through a number of manipulations, including counterstereotypic imagery (Blair, Ma, & Lenton, 2001) and exposure to positive group members (Dasgupta & Greenwald, 2001). In contrast, however, other research suggests that implicit attitudes are difficult to change even when detailed information in favor of a revised attitude is presented (e.g., Gregg et al., 2006; Rydell et al., 2006).

The present work yielded results that diverge from previous work. For example, in Gregg et al.'s (2006) work, participants who learned extensive information contradicting their initial beliefs about two groups failed to revise their implicit attitudes toward those groups. Several points differentiate the present research from that of Gregg et al. Most importantly, participants in the review evidence condition of Experiment 2 were not given new behavioral information, but were asked to review the same information they had read earlier—but this time with the knowledge that the target belonged to a different group than earlier believed. The fact that participants viewed the same information a second time, but with a new category driving their interpretation of it, makes it likely that participants' initial interpretations were tagged as false (Petty et al., 2006). As a result, when participants' impressions were later assessed, they would be unlikely to retrieve

 Au: All 3 →
 citations not in
 refs.

their original impressions but would rather respond on the basis of new, revised impressions.

The PAST model does not directly address this spectrum of situations where one set of beliefs may conflict with another. The divergent findings described above may appear to be contradictory on the surface, but may be understood in terms of Gawronski and Bodenhausen's (2006) APE model. According to the APE model, changes in implicit beliefs may occur as the result of new associations being learnt *or* as the result of a shift in which a subcomponent of a complex representation is currently most accessible. Findings from research reported by Rydell et al. (2006) are consistent with the former scenario, in that the introduction of new associations led to changes in implicit attitudes toward a target. In contrast, the latter possibility may account for results reported by Dasgupta and Greenwald (2001) and Blair et al. (2001) in which increasing the salience of distinct aspects of a complex category representation may have shifted the component of the representation that was temporarily most accessible.

How might the APE model account for the present results? Like those in previous studies, participants in the present experiment formed implicit stereotype-based associations which were not easily changed by simply manipulating participants' explicit beliefs about the applicability of the stereotype. Yet, participants' implicit beliefs about the target *did* change when they were able to reassess the behavioral evidence on which they had based those beliefs. On the surface, this finding may seem difficult to reconcile with previous research and with the APE model. However, the methods used in the present experiment differ in at least one important way from those used in previous research (Gregg et al., 2006; Rydell et al., 2006). Specifically, participants in previous studies have typically been exposed to *new* information that contradicted or counteracted the information they had initially learned. In contrast, participants in the research reported here were asked to reconsider the *same* information. The only "new" aspect of the information was the inference participants drew from it. Research from the person memory tradition (e.g., Winter & Uleman, 1984) suggests that people spontaneously associate behaviors with the traits that they reflect. However, the specific trait that is inferred from a given behavior may vary depending on the context (e.g., the group to which the actor belongs). Thus participants may have formed new associations with the behaviors they read, which may have then led them to overwrite previously formed associations with new associations which would then be applied to the target.

CONCLUSIONS

The research reported here extends prior work on the factors that lead to changes in implicit beliefs by demonstrating the critical importance of elaboration in determining the strength of both initial and revised beliefs, and hence the extent to which initial beliefs dictate implicit responses. The present study also presents an initial exploration of how stereotypic beliefs of an individual may persist in influencing responses even when those stereotypes have been discarded as invalid. Hence, these findings may have important implications for a variety of circumstances in which people change their beliefs about a target's membership in a social group.

type: refs should
be in 2 columns

REFERENCES

- Ashburn-Nardo, L., Voils, C. I., & Monteith, M. J. (2001). Implicit associations as the seeds of intergroup bias: How easily do they take root? *Journal of Personality and Social Psychology, 81*, 789-799.
- Blair, I. V. (2002). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review, 6*, 242-261.
- Blair, I. V., Ma, J. E., & Lenton, A. P. (2001). Imagining stereotypes away: The moderation of implicit stereotypes through mental imagery. *Journal of Personality and Social Psychology, 81*, 828-841.
- Dasgupta, N., & Asgari, S. (2004). Seeing is believing: Exposure to counterstereotypic women leaders and its effect on automatic gender stereotyping. *Journal of Experimental Social Psychology, 40*, 642-658.
- 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. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56*, 5-18. #
- Evans, J. St. B. T., & Over, D. E. (1996). *Reasoning and rationality*. Hove, UK: Psychology Press. # x 4
- Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin, 132*(5), 692-731.
- Golding, J. M., Fowler, S. B., Long, D. L., & Latta, A. (1990). Instructions to disregard potentially useful information: The effects of pragmatics on evaluative judgments and recall. *Journal of Memory and Language, 29*, 212-227.
- Greenwald, A. G., & Farnham, S. D. (2000). Using the Implicit Association Test (IAT) to measure self-esteem and self-concept. *Journal of Personality and Social Psychology, 79*, 1022-1038.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test (IAT). *Journal of Personality and Social Psychology, 74*, 1464-1480.
- Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test (IAT): I. An improved scoring algorithm. *Journal of Personality and Social Psychology, 85*, 197-216.
- 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.
- Kunda, Z., Davies, P. G., Adams, B. D., Spencer, S. J. (2002). The dynamic time course of stereotype activation: Activation, dissipation, and resurrection. *Journal of Personality and Social Psychology, 82*, 283-299. # x 3
- Nosek, B. A., & Banaji, M. R. (2001). The go/no-go association task. *Social Cognition, 19*, 625-666.
- Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math = male, me = female, therefore math ≠ me. *Journal of Personality and Social Psychology, 83*, 33-59.
- Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology, 19*, 123-205.
- Petty, R. E., Tormala, Z. L., Brinol, 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.
- Rydell, R. J., McConnell, A. R., Mackie, D. M., & Strain, L. M. (2006). Of two minds: Forming and changing valence inconsistent implicit and explicit attitudes. *Psychological Science, 17*, 954-958.
- Slooman, 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.
- Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review, 8*, 220-247.

- Wilson, T. D., Lindsey, S., & Schooler, T. (2000). A model of dual attitudes. *Psychological Review*, 107, 101-126.
- Winter, L., & Uleman, J. S. (1984). When are social judgments made? Evidence for the spontaneity of trait inferences. *Journal of Personality and Social Psychology*, 47, 237-252.
- Wyer, R.S., & Unverzagt, W.H. (1985). Effects of instructions to disregard information on its subsequent recall and use in making judgments. *Journal of Personality and Social Psychology*, 48, 533-549.

x 2