European Journal of Social Psychology
Eur. J. Soc. Psychol. **33**, 573–589 (2003)
Published online 7 August 2003 in Wiley InterScience (www.interscience.wiley.com). **DOI**: 10.1002/ejsp.166

Implicit bias in impression formation: associations influence the construal of individuating information

BERTRAM GAWRONSKI*, DANIEL GESCHKE AND RAINER BANSE
Humboldt-Universität zu Berlin, Germany

Abstract

The present research investigated the influence of group-related evaluative associations on the process of impression formation. In particular, we expected the impact of a target's category membership on the construal of ambiguous behaviour to be moderated by perceivers' evaluative associations related to the target category. Associative strength was further expected to have an indirect effect on dispositional inference, mediated by its impact on behaviour identification. Results support both of these assumptions. Moreover, the influence of evaluative associations on impression formation was not moderated by perceivers' motivation to control prejudiced reactions. Rather, motivation to control moderated only the relation between evaluative associations and the explicit endorsement of prejudiced beliefs about the target group in general, such that explicit prejudice endorsement was correlated with evaluative associations only for perceivers low, but not for those high in motivation to control. Implications for prejudice control are discussed. Copyright © 2003 John Wiley & Sons, Ltd.

When two individuals engage in the same behaviour, we often do not perceive the same act. For example, one and the same ambiguous shove may be judged as an aggressive act when the actor is black, but as a jovial nudge when the actor is white. Such biasing effects of category cues on the construal of individuating information have been demonstrated for a variety of social categories and behaviours (e.g. Darley & Gross, 1983; Duncan, 1976; Dunning & Sherman, 1997; Kunda & Sherman-Williams, 1993; Sagar & Schofield, 1980), and seem to be particularly pronounced for ambiguous behaviours (Dunning & Sherman, 1997; Kunda & Sherman-Williams, 1993).

Drawing on Fazio's (1990) MODE model of attitudes, these biasing effects may be explained by the automatic activation of evaluative or stereotypic concepts associated with the social category of the actor (see also Fazio, 1995). According to Fazio, automatically activated evaluative associations often bias the spontaneous construal of incoming stimulus information in an assimilative manner. Consistent with this assumption, Fazio and Williams (1986), for example, found that judgments of the performance of political candidates depended on perceivers' personal attitudes, such that the performance of personally favoured candidates was evaluated more positively than the performance of a

*Correspondence to: Dr B. Gawronski, Department of Psychology, Northwestern University, 2029 Sheridan Road, Evanston, IL 60208-2710, USA. E-mail: gawronski@northwestern.edu

Contract/grant sponsor: Deutsche Forschungsgemeinschaft (Ba 1603/1-1).

Received 26 March 2002
Accepted 2 March 2003

competing candidate. Most importantly, this biasing effect was moderated by the strength of the corresponding evaluative associations, such that it was more pronounced when evaluative associations were strong than when they were weak. Similar effects were obtained by Houston and Fazio (1989) who demonstrated that the evaluation of an article toward capital punishment depended on the judges' personal attitudes toward the issue in question. However, as with Fazio and Williams' (1986) results this effect was moderated by the chronic (Experiment 1) or situational activation (Experiment 2) of the corresponding evaluative associations (see also Schuette & Fazio, 1995).

Applied to the present issue of category effects on the construal of ambiguous individuating information, one could argue that the spontaneous categorization of an actor automatically activates constructs stereotypically associated with his or her social category (Devine, 1989). These activated constructs may then prime a particular construal of the behaviour, which in turn affects the general impression of the target (e.g. Higgins, Rholes, & Jones, 1977; Srull & Wyer, 1979; see also Trope, 1986). In other words, context effects of category cues on the construal of ambiguous behaviour might be due to the automatic activation of stereotypic or evaluative associations related to the actor's category, which then primes a particular construal of the observed behaviour.

The main goal of the present research was to investigate the moderating role of associative strength for the impact of category cues on the construal of ambiguous behaviour. Drawing on previous evidence for biasing effects of evaluative associations (e.g. Fazio & Williams, 1986; Houston & Fazio, 1989; Schuette & Fazio, 1995), it is hypothesized that context effects of category cues on behaviour identification depend on the strength of category-related evaluative associations. Specifically, it is assumed that the impact of category information on the construal of ambiguous behaviour is stronger when category-related associations are strong than when they are weak. Moreover, category-related associations were expected to indirectly bias dispositional inferences about an actor, mediated by their impact on behaviour identification (see Trope, 1986).

DIRECT AND INDIRECT EFFECTS OF ASSOCIATIONS

An important aspect of the present hypotheses concerns the way of how associations bias the process of impression formation. Based on the present reconstruction, it is assumed that associations affect the subjective construal of the observed behaviour, and thus dispositional inferences about the actor (Kunda & Thagard, 1996; Trope, 1986). Another possibility, however, is that perceivers take the ambiguous behaviour as it is and use their category-related associations as an independent information for a dispositional judgment about the target (Brewer, 1988; Fiske & Neuberg, 1990). Even though both of these processes are theoretically possible, previous evidence on this issue clearly supports the former assumption. Srull and Wyer (1980), for example, found that trait priming affects impression formation (e.g. Higgins et al., 1977; Srull & Wyer, 1979) only when traits are activated before behavioural information is encoded, but not when traits are activated afterwards. If activated trait constructs actually have a direct effect on dispositional inference, order of trait priming and encoding of the behaviour should have no effect on dispositional attributions. This conclusion of an indirect effect is also supported by Newman and Uleman (1990) who obtained trait priming effects on the interpretation of ambiguous behaviour in a spontaneous trait inference paradigm (see Uleman, Newman, & Moskowitz, 1996, for a review). In contrast to Srull and Wyer (1980), however, Newman and Uleman (1990) used an implicit measure of spontaneous trait inference, suggesting that perceivers are not aware of the obtained influence of trait primes. Similar results were found by Otten and Moskowitz (2000) who demonstrated that the spontaneous interpretation of ambiguous behavioural episodes varies as a function of the target's group membership (in-group vs. out-group) manipulated

by a minimal group paradigm. Taken together, these results suggest that activated traits or evaluations affect impression formation in an indirect way, i.e. mediated by their impact on behaviour identification. However, even though many researchers seem to agree with this conclusion (e.g. Dunning & Sherman, 1997; Fazio, 1990; Kunda & Sherman-Williams, 1993), most studies only used trait attribution measures, thus allowing for both of these interpretations (e.g. Devine, 1989; Lepore & Brown, 1997). Hence, a second goal of the present study is to disentangle the proposed influence of evaluative associations on behaviour identification from a direct impact on dispositional attributions.

MOTIVATION TO CONTROL PREJUDICED REACTIONS

A third question addressed in the present study is whether perceivers adjust a biased construal of ambiguous behaviour to explicit egalitarian goals. Specifically, one could suspect that perceivers with a high motivation to control prejudiced reactions (Dunton & Fazio, 1997; Plant & Devine, 1998) explicitly attempt to correct their judgments for the biasing potential of category-related associations. Such explicit debiasing attempts, however, should emerge only when perceivers are actually aware of such an influence (Strack & Hannover, 1996; Wegener & Petty, 1997). If they are not aware of the biasing influence, but take their subjective construal as an objective given, motivation to control prejudiced reactions seems rather unlikely to have a moderating effect on the proposed impact of evaluative associations. Preliminary evidence for the latter assumption can be drawn from a study by Trope and Alfieri (1997, Experiment 2) who found that even an explicit invalidation of context information does not reduce its biasing effect on the construal of behavioural information. Drawing on these results, a third objective of the present study was to test whether the proposed influence of evaluative associations on impression formation is moderated by individual differences in the motivation to control prejudiced reactions.

PREVIOUS EVIDENCE FOR BIASING EFFECTS OF ASSOCIATIONS

Preliminary evidence for the assumption that category-related associations moderate the impact of category information on the construal of ambiguous individuating information can be found in a study by Lepore and Brown (1997, Experiment 2). In this study, white participants were primed either to the social category of black people or to the category of white people. Immediately after the priming task, participants were asked to read a number of behavioural descriptions about an ethnically unspecified target and to form an impression of the actor. Results indicate that impressions by highly prejudiced participants were more negative when they were primed to the category of black people than when they were primed to the category of white people. Participants with a low prejudice level, in contrast, were unaffected by the category priming. Drawing on these results, Lepore and Brown (1997) argued that high and low prejudiced individuals may differ with respect to their evaluative associations related to the category of black people. Whereas low prejudice individuals may have weak or neutral associations with respect to black people, high prejudice individuals may have strong negative associations. These associations, in turn, are assumed to be activated automatically upon the encounter of a category-related stimulus, which in turn leads to a biased impression of the target.

Even though Lepore and Brown's (1997) results offer initial evidence for the present hypothesis that the construal of ambiguous behavioural information is biased by evaluative associations, their study offers no conclusive evidence for this assumption. First, Lepore and Brown (1997) used only a trait inference measure to assess impression formation. Such kind of measures, however, are generally

ambiguous as to whether the obtained effects are due to a biased interpretation of the target's behaviour, or to a direct effect on trait attribution (Trope, 1986). Second, Lepore and Brown (1997) assessed explicit prejudice endorsement, rather than associative strength. Recent results, however, suggest that evaluative associations and explicit prejudice endorsement are not necessarily related to one another. Rather, category-related evaluative associations often seem to dissociate from the explicit endorsement of prejudiced beliefs (e.g. Fazio, Jackson, Dunton, & Williams, 1995; Greenwald, McGhee, & Schwartz, 1998; Wittenbrink, Judd, & Park, 1997; for reviews see Blair, 2001; Dovidio, Kawakami, & Beach, 2001). Hence, it is not clear whether Lepore and Brown's (1997) results are actually due to a differential activation of negative associations or to a more explicit process involved.

In order to solve these interpretational ambiguities, the present study included two distinct measures for behaviour identification and dispositional inference. Moreover, strength of category-related evaluative associations was assessed with Greenwald et al.'s (1998) Implicit Association Test (IAT), rather than with an explicit measure of prejudice endorsement. Based on the considerations outlined above, we expected context effects of category cues on the construal of ambiguous behaviour to be moderated by the strength of category-related evaluative associations assessed with an IAT. Moreover, evaluative associations were predicted to have an indirect effect on the attribution of stable dispositions, mediated by their impact on behaviour identification. Finally, we assumed that perceivers are unaware of the biasing influence of their associations, thus implying that individual differences in the motivation to control prejudiced reactions (Dunton & Fazio, 1997) have no moderating effect on the predicted effect of evaluative associations. A moderating influence of motivation to control was expected only for cases in which perceivers are explicitly aware of the biasing influence of their associations, such as can be assumed for the explicit endorsement of prejudiced beliefs about the target group in general (e.g. Banse & Gawronski, 2003; Banse, Seise, & Zerbes, 2001; Dunton & Fazio, 1997; Fazio et al., 1995).

METHOD

Overview

German participants read a written description about an ambiguously behaving male target. Ethnic origin of the target was indicated by a portrait photo of either a typically Turkish- or German-looking young man. After reading the description, participants were asked to evaluate the behaviour of the target (behaviour identification) and to predict his behaviour in a number of different situations (dispositional inference). Finally, participants' explicit prejudiced beliefs about Turkish people, their motivation to control prejudiced reactions, and strength of negative associations related to Turkish as compared to German people were assessed.

Participants and Design

A total of 70 participants (37 female) drawn from a volunteer pool took part in a study on interpersonal perception in return for the payment of 10 Deutsche Mark ($\approx$ 5 € or US-$ 5). Participants were randomly assigned to one of the two experimental conditions (i.e. Turkish vs. German target). Together with the two quasi-experimental factors of associative strength and motivation to control prejudiced reactions, the study consisted of a 2 (ethnic origin: Turkish vs. German) $\times$ 2 (negative associations: weak vs. strong) $\times$ 2 (motivation to control: low vs. high) factorial design. Experimental sessions were run individually. Data from one participant who terminated the experimental session had to be excluded from analyses.

Procedure

When participants arrived they were welcomed and informed that they were taking part in a study on interpersonal perception. The experimenter explained that they were to read a short story about an evening in the life of a young male target named A, and that their task was to form an impression of this person. They were then handed a booklet with a black-and-white portrait photo of either a typically Turkish- or German-looking young man, and a short story about this person. Photographs were taken from the Psychological Image Collection at Stirling (University of Stirling Psychology Department, 2000). Photos were selected by pretests. The story described an evening the target spent with some friends at a disco. The target's behaviour was held ambiguous with respect to a positive or a negative interpretation (see Srull & Wyer, 1979). In particular, the critical ambiguous behavioural episodes were: (a) approaching a woman on the dance floor and asking for her telephone number, (b) getting stuck in a discussion with his friends about 'how women really are', and (c) driving home alone afterwards. When participants had read the story, they were asked to evaluate the target's behaviour (behaviour identification), and to predict his behaviour in several hypothetical situations that could evoke either positive or negative behaviour (dispositional inference). Afterwards, participants were asked to fill in a manipulation check on the perceived ethnic origin of the target, Pettigrew and Meertens' (1995) Subtle and Blatant Prejudice Scales, and a German adaptation of Dunton and Fazio's (1997) Motivation to Control Prejudiced Reactions Scale (Banse & Gawronski, 2003). After completion of the questionnaire, participants were administered an IAT (Greenwald et al., 1998) to assess the strength of negative associations toward Turkish as compared to German people. Finally, participants were probed for suspicion, debriefed, and thanked for participation.[1]

Measures

Behaviour Identification

In order to assess participants' subjective construal of the target's behaviour, they were asked to judge his behaviour with respect to eight behavioural dimensions: objectionable (unangenehm), brash (aufdringlich), cheeky (frech), obtrusive (penetrant), insensitive (taktlos), obstinate (stur), arrogant (arrogant), bigheaded (eingebildet). Ratings of the target's behaviour were assessed with rating scales ranging from 1 (= not true) to 5 (= true).

Dispositional Inference

To assess participants' dispositional inferences about the target they were asked to predict the target's behaviour in nine hypothetical situations that could elicit either positive or negative behaviour (see Appendix). The subjective likelihood of negative behaviour had to be rated on scales ranging from 1 (= very unlikely) to 5 (= very likely).

[1]Drawing on recent evidence for context effects on the IAT (e.g. Blair, Ma, & Lenton, 2001; Dasgupta & Greenwald, 2001; Wittenbrink, Judd, & Park, 2001) one may criticize that evaluative associations were assessed after rather than before the impression formation task. More precisely, IAT-scores could have been affected by the experimental manipulations, thus undermining an interpretation of the obtained results in terms of the present hypotheses. This assumption, however, was not confirmed for the present data. Specifically, a multivariate analyses of variance (MANOVA) did not reveal any significant effect of the experimental manipulations on evaluative evaluations, explicit prejudice endorsement, or motivation to control prejudiced reactions (see also Gawronski, Ehrenberg, Banse, Zukova, & Klauer, 2003; Lepore & Brown, 1997).

Manipulation Checks

In order to assure that the targets were actually perceived as Turkish or German, the target had to be rated with respect to his ethnic origin (i.e. German, Turkish) on two scales ranging from 1 (very unlikely) to 5 (= very likely).

Motivation to Control Prejudiced Reactions

Participants' idiosyncratic motivation to control prejudiced reactions was assessed with a German adaptation of Dunton and Fazio's (1997) Motivation to Control Prejudiced Reactions Scale (Banse & Gawronski, 2003). Five-point scales were used for assessment.

Explicit Prejudice Endorsement

Endorsement of explicit prejudiced beliefs about Turkish people was assessed with a German version of Pettigrew and Meertens' (1995) Subtle and Blatant Prejudice Scales (Zick, 1997) using 5-point scales for assessment.

Evaluative Associations

To assess the strength of negative associations with respect to Turkish as compared to German people, an adaptation of Greenwald et al.'s (1998) IAT was used. The IAT was run on a 486 IBM-compatible laptop using the software Experimental Run Time System ERTS (Beringer, 1994). Following Greenwald et al. (1998), the IAT consisted of five blocks (see Table 1). In the initial target-concept discrimination task (Block 1), portrait photographs of nine Turkish and nine German individuals had to be assigned to the categories 'Turkish' or 'German', respectively. Participants were asked to press a left-hand key ('A') when a German face appeared on the screen, and a right-hand key ('L') in the case of a Turkish face. In the attribute discrimination task (Block 2), nine positive and nine negative nouns adapted from Schwibbe, Räder, Schwibbe, Borchardt, and Geiken-Pophanken (1994) were presented and had to be classified according to the categories negative (left-hand key) and positive (right-hand key). In the initial combined task (Block 3), target and attribute discrimination trials were combined in a prejudice-inconsistent manner. Participants had to press the left-hand key when either a German face or a negative noun was presented, and the right-hand key when a Turkish face or a positive noun was

Table 1. Task sequences of the IAT used to assess strength of negative associations related to Turkish as compared to German people

			Response key assignment	
Block	Task	Trials	Left key	Right key
1	Initial target discrimination	36	German	Turkish
2	Attribute discrimination	36	Negative	Positive
3	Initial combined task	108	Negative, German	Positive, Turkish
4	Reversed target discrimination	36	Turkish	German
5	Reversed combined task	108	Negative, Turkish	Positive, German

presented. In the reversed target-concept discrimination task (Block 4), the initial target-concept discrimination was repeated with a switch of the categorization keys. The reversed combined task (Block 5) again combined the two individual tasks, now in a prejudice-consistent manner. Participants had to press the left-hand key when either a Turkish face or a negative noun was presented, and the right-hand key when a German face or a positive noun was presented. Each block started with a brief instruction for the following task and a request to respond as fast as possible even if this would lead to some errors. The three discrimination tasks (Blocks 1, 2, and 4) consisted of a total of 36 trials, respectively. The two combined tasks (Blocks 3 and 5) each comprised 108 trials (54 faces, 54 nouns), with targets and attributes being presented in an alternating order. A predetermined random order of trials was used for all participants. The response-stimulus interval following correct responses was 250 ms. Wrong responses were indicated with the word 'FEHLER!' (German for 'Error!') appearing for 1000 ms below the centre of the screen.[2]

RESULTS

Preliminary Analyses

Evaluative Associations

Following Greenwald et al. (1998), response latencies lower than 300 ms were recoded to 300 ms, and latencies higher than 3000 ms were recoded to 3000 ms. Error trials were excluded from analyses. Individual IAT-scores were calculated by first log-transforming response latencies and then subtracting the mean response time of the reversed combined task (Block 5) from the mean latency of the initial combined task (Block 3). This score was interpreted as an index for the strength of negative associations towards Turkish as compared to German people. Raw IAT scores ranged from -440.13 ms to $+655.86$ ms ($M = 30.88$, $SD = 211.28$). In order to estimate the reliability of the IAT (see Gawronski et al., 2003), the two combined blocks were each divided into three consecutive parts of equal length (i.e. 36 trials). The three thirds were then used to calculate three IAT-scores for each participant, which revealed an internal consistency of 0.90 (Cronbach's α). In order to obtain groups with weak and strong negative associations the sample was divided by a median-split of the log-transformed score ($MD = -0.03$).

Motivation to Control Prejudiced Reactions

Ratings obtained in the Motivation to Control Prejudiced Reactions Scale were merged into a single index by calculating their mean values (Cronbach's $\alpha = 0.74$). This index ranged from 2.63 to 4.81 ($M = 3.68$, $SD = 0.51$) and revealed a negative but non-significant correlation with the IAT-score (see Table 2). In order to obtain groups with high and low motivation to control prejudiced reactions the sample was divided by a median-split ($MD = 3.75$).

[2]One may criticize that we did not counterbalance prejudice-consistent and prejudice-inconsistent blocks in the IAT. Hence, the obtained IAT-scores could have been affected by order effects, thus undermining an interpretation of the so-called IAT-effect (Greenwald & Nosek, 2001). However, even though counterbalancing may be useful when the IAT is used as a dependent measure, it is actually less appropriate when the IAT is used as an independent measure. Specifically, due to the well established order effects (Greenwald et al., 1998) counterbalancing can produce two different distributions that are not comparable to one another, i.e. one and the same difference score may represent a high score in one distribution and a low score in the other. Hence, when IAT-scores are used as a predictor variable collapsing such incomparable distributions has the potential to attenuate any theoretically meaningful relation to other measures (see Gawronski, 2002, for a discussion).

Table 2. Correlations between strength of negative associations (implicit prejudice), explicit prejudice endorsement (explicit prejudice), and motivation to control prejudiced reactions

	1	2	3
1 Implicit prejudice	(0.90)	0.22	−0.16
2 Explicit prejudice		(0.87)	−0.26*
3 Motivation to control			(0.74)

Note: Cronbach's α estimates of internal consistency are in parentheses; $^*p < 0.05$.

Manipulation Checks

To assure that the targets were actually perceived as Turkish or German, ratings of the suspected ethnic origin were merged into a single index by recoding the ratings for a German origin. Hence, higher ratings indicate an attribution of a Turkish origin, and lower ratings an attribution of a German origin. This index was submitted to a 2 (ethnic origin: Turkish vs. German) × 2 (negative associations: weak vs. strong) × 2 (motivation to control: low vs. high) analysis of variance (ANOVA), revealing a significant main effect of the target's ethnic origin, $F(1, 61) = 62.67$, $p < 0.001$. Consistent with the intended manipulation, the proposed German target was rated less likely to be Turkish rather than German ($M = 1.69$) and the proposed Turkish target was rated more likely to be Turkish rather than German ($M = 3.31$). No other main or interaction effect reached statistical significance.

Behaviour Identification

To test the predicted moderator effect of evaluative associations on behaviour identification, ratings of the target's behaviour were merged into a single index by calculating their mean values (Cronbach's $\alpha = 0.92$). This index was submitted to a 2 (ethnic origin: Turkish vs. German) × 2 (negative associations: weak vs. strong) × 2 (motivation to control: low vs. high) analysis of variance (ANOVA). Replicating previous evidence for category effects on behaviour identification (e.g. Darley & Gross, 1983; Duncan, 1976; Dunning & Sherman, 1997; Kunda & Sherman-Williams, 1993; Sagar & Schofield, 1980), this analysis revealed a significant main effect of the target's ethnic origin, $F(1, 61) = 9.83$, $p < 0.01$, indicating that the described behaviour was rated less negative when the target was German ($M = 2.00$) than when he was Turkish ($M = 2.71$). This main effect was qualified by the expected two-way interaction of ethnic origin and associative strength, $F(1, 61) = 9.27$, $p < 0.01$. Specifically, participants with strong negative associations towards Turkish people rated the behaviour as more negative when the target was Turkish than when he was German ($M_{\text{Turkish}} = 3.13$, $M_{\text{German}} = 1.72$), $t(32) = 3.87$, $p < 0.001$. However, when associations were weak, ratings of the behaviour were unaffected by the target's ethnic origin ($M_{\text{Turkish}} = 2.21$, $M_{\text{German}} = 2.22$), $t(33) = -0.22$, *ns*. Also consistent with the present predictions, this interaction was not qualified by participants' motivation to control prejudiced reactions, $F(1, 61) = 0.03$, *ns* (see Table 3). No other main or interaction effect reached statistical significance.

Dispositional Inference

To test the impact of evaluative associations on participants' dispositional inferences, likelihood ratings of negative behaviour in hypothetical situations were merged into a single index by calculating mean values (Cronbach's $\alpha = 0.72$). This index was submitted to a 2 (ethnic origin: Turkish vs.

Table 3. Mean values of behaviour identification and dispositional inference as a function of the target's ethnic origin, strength of negative associations, and motivation to control prejudiced reactions

Ethnicity	Low motivation to control		High motivation to control	
	Weak associations	Strong associations	Weak associations	Strong associations
Behaviour identification				
Turkish	2.37	3.06	2.11	3.29
German	2.19	1.47	2.25	1.85
Dispositional inference				
Turkish	3.22	3.26	2.92	3.48
German	3.02	2.44	2.90	2.76

Note: Higher values indicate more negative judgments.

German) × 2 (negative associations: weak vs. strong) × 2 (motivation to control: low vs. high) analysis of variance (ANOVA), revealing a significant main effect of the target's ethnic origin, $F(1, 61) = 8.30$, $p < 0.01$. In particular, the Turkish target was rated more likely to behave negatively than the German target ($M_{\text{Turkish}} = 3.20$, $M_{\text{German}} = 2.83$). This main effect was qualified by a significant two-way interaction of ethnic origin and associative strength, $F(1, 61) = 4.66$, $p < 0.05$. As expected, participants with strong negative associations towards Turkish people rated the target more likely to behave negatively when he was Turkish than when he was German ($M_{\text{Turkish}} = 3.33$, $M_{\text{German}} = 2.65$), $t(32) = 3.18$, $p < 0.01$. In contrast, participants with weak associations rated the Turkish and the German target approximately equal in the likelihood for negative behaviour ($M_{\text{Turkish}} = 3.03$, $M_{\text{German}} = 2.97$), $t(33) = 0.33$, *ns*. Also consistent with the present predictions, this interaction was not qualified by participants' motivation to control prejudiced reactions, $F(1, 61) = 0.02$, *ns* (see Table 3). No other main or interaction effect reached statistical significance.

Behaviour Identification as Mediator for Dispositional Inference

Evaluative associations were expected to have a direct effect on behaviour identification, but an indirect effect on dispositional inference which should be mediated by their impact on behaviour identification. In order to test this hypothesis, an analysis of covariance (ANCOVA) was conducted with ethnic origin, associative strength, and motivation to control as fixed factors, behaviour identification as covariate, and dispositional inference as dependent measure. If the obtained effect of associative strength on dispositional inference is mediated by behaviour identification, this analysis should reveal a significant effect of behaviour identification, whilst the interaction of associative strength and ethnic origin should fail to reach statistical significance. If, however, associative strength has a direct effect on dispositional inference, the interaction of associative strength and ethnic origin should still reach statistical significance even when it is controlled for behaviour identification. Consistent with the present predictions, results clearly support the former assumption. Specifically, the conducted ANCOVA revealed a highly significant effect of behaviour identification, $F(1, 60) = 20.13$, $p < 0.001$, with negative judgments of the behaviour being positively related to the inference of a negative disposition ($r = 0.59$, $p < 0.001$). The previously obtained interaction of associative strength and ethnic origin, however, failed to reach statistical significance after controlling for behaviour identification, $F(1, 60) = 0.46$, *ns*. A Sobel test indicated a significant mediation of behaviour identification, $z = 2.52$, $p < 0.05$. No other main or interaction effect reached statistical significance.

Explicit Prejudice Endorsement

Because perceivers often take a biased construal of behavioural information as an objective given (e.g. Trope & Alfieri, 1997), explicit attempts to correct for the influence of evaluative associations are rather unlikely. This assumption is clearly supported by the present data. However, participants can be expected to adjust their judgments to explicit egalitarian goals when they are aware of a biasing influence (Strack & Hannover, 1996; Wegener & Petty, 1997). This should be the case, for example, when participants are asked to report their explicit beliefs about the target group in general (e.g. Banse & Gawronski, 2003; Banse et al., 2001; Dunton & Fazio, 1997; Fazio et al., 1995). Drawing on these considerations, we expected motivation to control prejudiced reactions to moderate the relation between the strength of negative associations and the explicit endorsement of prejudiced beliefs. Specifically, whereas participants with a high motivation to control should exhibit a low level of explicit prejudice regardless of the strength of negative associations, participants with a low motivation to control should exhibit a high level of explicit prejudice when their associations are strongly negative, but a low level when negative associations are weak.

In order to test this assumption, ratings of subtle and blatant prejudice were merged into a single index of explicit prejudice against Turkish people (Cronbach's $\alpha = 0.87$).[3] This index ranged from 1.44 to 4.22 ($M = 2.66$, $SD = 0.59$) and exhibited a marginally significant positive correlation with the IAT-score of negative associations towards Turkish people and a significant negative correlation with motivation to control prejudiced reactions (see Table 2). With respect to impression formation, there were no significant effects of explicit prejudice endorsement on behaviour identification or dispositional inference. However, a 2 (negative associations: weak vs. strong) $\times$ 2 (motivation to control: low vs. high) analysis of variance (ANOVA) on explicit prejudice endorsement revealed a significant main effect of motivation to control, $F(1, 65) = 8.62$, $p < 0.01$, indicating that participants high in motivation to control exhibited a lower level of explicit prejudice ($M = 2.47$) than participants low in motivation to control ($M = 2.85$). Most importantly, this main effect was qualified by a marginally significant two-way interaction of associative strength and motivation to control, $F(1, 65) = 3.50$, $p = 0.07$. As expected, participants low in motivation to control exhibited a higher level of explicit prejudice when they had strong negative associations than when they had weak associations ($M_{\text{weak}} = 2.65$, $M_{\text{strong}} = 3.06$), $t(33) = 2.14$, $p < 0.05$. In contrast, participants high in motivation to control exhibited a low level of explicit prejudice regardless of the strength of negative associations ($M_{\text{weak}} = 2.51$, $M_{\text{strong}} = 2.42$), $t(32) = -0.47$, *ns*. Moreover, whereas participants with strong associations revealed a higher level of explicit prejudice when motivation to control was low than when it was high, $t(32) = 3.15$, $p < 0.01$, participants with weak associations exhibited an approximately equal level of explicit prejudice regardless of motivation to control, $t(33) = 0.82$, *ns*. This pattern is also reflected by correlational analyses, revealing a significant positive correlation between associative strength and explicit prejudice endorsement for participants with a low motivation to control prejudiced reactions ($r = 0.37$, $p < 0.05$), but a zero-correlation for participants with a high motivation to control ($r = -0.02$, *ns*). The difference between the two correlations is statistically significant, $z = 1.64$, $p < 0.05$ (one-sided).

DISCUSSION

The main goal of the present study was to test whether the impact of social category cues on the construal of ambiguous behavioural information (e.g. Darley & Gross, 1983; Duncan, 1976; Dunning

[3]Since subtle and blatant prejudice were highly correlated ($r = 0.61$, $p < 0.001$), the two measures were merged into a single index of explicit prejudice endorsement.

& Sherman, 1997; Kunda & Sherman-Williams, 1993; Sagar & Schofield, 1980) is moderated by the strength of category-related evaluative associations. The present data clearly support this assumption. Specifically, context effects of category cues on behaviour identification were limited to perceivers with strong category-related evaluative associations. For perceivers with weak or neutral associations, in contrast, behaviour identification was unaffected by category cues. Also consistent with the present predictions, this effect of assimilative behaviour identification further affected perceivers' dispositional inferences about the target. This mediation was indicated by (1) a significant interaction of associative strength and the target's ethnicity in the prediction of behaviour identification, (2) a significant interaction of associative strength and the target's ethnicity in the prediction of dispositional inference, (3) a significant relation between behaviour identification and dispositional inference, and (4) a reduction to non-significance of the interaction between associative strength and the target's ethnicity in the prediction of dispositional inference after controlling for behaviour identification.

Another interesting result is that the obtained influence of evaluative associations on impression formation was not qualified by participants' motivation to control prejudiced reactions. This was true for both behaviour identification and dispositional inference. Hence, it seems that perceivers are actually unaware of the biasing potential of their associations with respect to the construal of ambiguous behaviour. In contrast, they seem to take the biased construal of behavioural information as an objective given (Trope & Alfieri, 1997), rather than as a product of their idiosyncratic associations. A moderating influence of motivation to control prejudiced reactions was found only with respect to the relation between evaluative associations and the explicit endorsement of prejudiced beliefs about the target group in general. Specifically, participants high in motivation to control exhibited a low level of explicit prejudice regardless of the strength of negative associations. In contrast, participants low in motivation to control showed a higher level of explicit prejudice when negative associations were strong than when they were weak.

By offering evidence for a biasing influence of evaluative associations on the construal of ambiguous behaviour, the present results extend the contribution of previous studies demonstrating effects of stereotypic associations on category-based and individuating impression formation (see Brewer, 1988; Fiske & Neuberg, 1990). Gawronski et al. (2003), for example, demonstrated that a target's category membership affected the ascription of stereotypical traits only when perceivers had strong stereotypical associations. For perceivers with weak associations, in contrast, impressions were affected only by individuating, but not by category information. Moreover, employing Klauer and Wegener's (1998) multinomial model of the 'Who said what?' paradigm (Taylor, Fiske, Etcoff, & Ruderman, 1978), these effects were demonstrated to be due to increased stereotyping and decreased individuation for perceivers with strong stereotypic associations, rather than to individual differences in the likelihood of social categorization. The present results extend these findings by demonstrating an additional way of how impression formation can be biased by category-related associations.

Possible Criticism

Even though the present results seem to unambiguously support our predictions, one may criticize that the obtained mediation effect of behaviour identification on dispositional inference could be spurious since the measures used in the present experiment might assess the same construct: trait attributions to the actor. Moreover, the present order of assessing behaviour identification and dispositional inference could have promoted the obtained mediation, such that participants strongly relied on their previous judgments of the target's behaviour when later predicting negative behaviour in the described scenarios. Even though we cannot rule out the possibility of a promoting order effect, we strongly believe that behavioural judgments and behavioural predictions refer to independent constructs that do

not correspond to one another. Using similar measures for behaviour identification and dispositional inference, Gawronski, Alshut, Grafe, Nespethal, Ruhmland, and Schulz (2002), for example, found differential effects of situational context information on behavioural judgments and behavioural predictions. Whereas information about the situational context led to assimilation effects in judgments of the target's behaviour, the same information produced contrast effects with respect to behavioural predictions (see also Trope, Cohen, & Alfieri, 1991; Trope, Cohen, & Maoz, 1988). This result indicates (a) that effects on behavioural judgments and behavioural predictions do not necessarily correspond to one another, and thus (b) that behavioural judgments and behavioural predictions are not just different measures for the same construct (see also Jackson, Sullivan, & Hodge, 1993).

Another objection may be that our treatment of the present IAT-data implies a comparison between perceivers with positive versus negative associations related to Turkish people, rather than a comparison between perceivers with strong versus weak negative associations. Specifically, the IAT-scores obtained in the present experiment ranged from −440 to +656 ms, thus suggesting that participants below the median have positive rather than weak negative or neutral associations. Even though we cannot unambiguously rule out this possibility, there are a number of arguments that support the present interpretation in terms of weak or neutral associations. First of all, IAT-scores are generally affected by the order of consistent and inconsistent blocks (Greenwald et al., 1998). Thus, an absolute interpretation of IAT-scores seems generally unwarranted. Most importantly, using an inconsistent-consistent order, such as employed in the present study, usually moves distributions of IAT-scores into the direction of lower scores (i.e. it reduces the so-called IAT-effect). Hence, the obtained negative values do not necessarily reflect positive associations toward Turkish people. Second, and directly related to this point, it seems generally difficult to determine an absolute value for neutral associations in the IAT (Greenwald & Nosek, 2001). Steffens and Plewe (2001), for example, found that IAT-scores can be strongly affected by the items presented in the IAT. Hence, distributions of IAT-scores can vary as a function of the stimulus material. This, however, also implies that negative IAT-scores do not necessarily reflect positive associations. Third, using a tripartite instead of a median split of the present IAT-data we found continuously increasing effect sizes of the target's ethnicity as a function of increasing associative strength. With respect to dispositional inference, for example, effect sizes (Cohen's *d*) increased from 0.14 (low IAT-scores), over 0.63 (medium IAT-scores) up to 0.96 (high IAT-scores). If negative IAT-scores actually reflect positive associations and IAT-scores around zero neutral associations, participants with medium IAT-scores should exhibit no effect of the target's ethnicity. If, however, negative scores reflect neutral associations and scores around zero weak negative associations, one can expect a continuously increasing effect of the target's ethnicity as a function of associative strength, as was obtained in the present experiment. Fourth, if the obtained negative IAT-scores actually represent positive associations, this would imply that impression formation is biased only by negative, but not by positive associations. This, however, would be in contrast to the results of previous studies which generally obtained corresponding effects for positive and negative associations (e.g. Fazio & Williams, 1986; Houston & Fazio, 1989; Schuette & Fazio, 1995). Hence, it seems more likely that negative IAT-scores reflect neutral rather than positive associations, and scores around zero weak negative rather than neutral associations. Finally, the IAT assesses only relative rather than absolute evaluations (Greenwald & Nosek, 2001). IAT-scores in the present experiment, for example, reflect a relative preference for Germans in contrast to Turkish people. Hence, an absolute interpretation of the obtained IAT-scores would imply that some of our German participants had a stronger preference for the out-group as compared to their in-group. This, however, seems rather unlikely in the light of previous evidence for the generality of in-group favouritism (Brewer & Brown, 1998).

In sum, it seems that an absolute interpretation of IAT-scores is generally not warranted. Hence, the obtained negative IAT-scores do not necessarily reflect positive associations with respect to Turkish

people. In contrast, these negative scores may be more likely to reflect an equal evaluation of German and Turkish people rather than a preference for Turkish over German people. This conclusion is consistent with the proposed interpretation that biasing effects of category cues are more pronounced when category-related associations are strong than when they are weak.

Implications for Prejudice Control

From a general perspective, the present results suggest that prejudice-related associations may have a self-confirming character, which makes them quite difficult to change. In particular, prejudice-related associations seem to lead perceivers to 'see' what they already have in their mind (Bruner, 1957). Hence, the biased construal of ambiguous behaviour performed by a stereotyped target may further bolster the associative links between the stereotyped group and negative behaviours, thus resulting in a kind of vicious circle.

Another important implication of the present results concerns attempts to avoid prejudiced reactions by thoroughly attending to individuating information (e.g. Brewer, 1988; Fiske & Neuberg, 1990). Specifically, the present results suggest that this strategy may be less than optimal when perceivers have strong negative associations and behavioural evidence is ambiguous. In this case, perceivers' associations may bias the construal of the available individuating information, thus leading them to see what they already have in their mind. Hence, even when perceivers are highly motivated to avoid prejudiced reactions, and thus attempt to base their reactions on individuating information, their reactions may still be prejudiced due to the biased construal of this information.

Even though the considerations thus far may suggest a rather pessimistic conclusion, the present results also imply some positive aspects. First of all, the impact of social category cues on the construal of individuating information seems to be not as universal as previous results may suggest (e.g. Devine, 1989; Dunning & Sherman, 1997). In contrast, it seems that how we perceive one and the same behaviour depends not only on who is acting, but also on our chronic associations. Hence, a very fundamental way to escape the vicious circle described above would be a modification of these associations. Recent results suggest that such attempts can be quite effective. Rudman, Ashmore, and Gary (2001), for example, found that negative associations with respect to black people can be modified through diversity education. Similar results were found by Kawakami, Dovidio, Moll, Hermsen, and Russin (2000) using a stereotype negation training. Furthermore, Moskowitz and his colleagues have demonstrated that chronic egalitarian goals have the potential to inhibit the automatic activation of negative or stereotypical associations (e.g. Moskowitz, Gollwitzer, Wasel, & Schaal, 1999; Moskowitz, Salomon, & Taylor, 2000), which can be assumed to have the same effect on impression formation as proposed for a modification of associations.

Second, the obtained null effect of motivation to control prejudiced reactions with respect to impression formation particularly concerns explicit attempts for judgmental correction. In contrast to such *explicit* correctional goals, however, Bargh, Gollwitzer, Lee-Chai, Barndollar, and Trötschel (2001) recently demonstrated that correctional goals can also operate outside of awareness. Hence, it is possible that such *implicit* correctional goals can reduce the biasing influence of evaluative associations on impression formation. Future research may help to clarify whether preconscious egalitarian goals can actually help to overcome the obtained effects of evaluative associations on the construal of ambiguous behaviour.

Finally, the present results do not imply that perceivers generally cannot be aware of the biasing influence of their associations. Rather, this could be an insightful lesson for some meta-cognitive advice. In other words, teaching lay people about the functioning of associations may increase their awareness of the biasing potential of their personal associations, which in turn may evoke explicit correctional goals for the particular kind of bias reported in the present study.

Conclusion

The main goal of the present study was to test whether the impact of social category cues on the construal of ambiguous individuating information is moderated by perceivers' category-related associations. Consistent with this assumption, the biasing influence of social category cues on the construal of ambiguous behaviour was demonstrated to increase as a function of the strength of category-related associations. Moreover, evaluative associations were demonstrated to indirectly bias dispositional inferences about the target, mediated by their impact on behaviour identification. Most interestingly, these effects were not moderated by perceivers' motivation to control prejudiced reactions. Rather, motivation to control prejudiced reactions moderated only the relation between evaluative associations and the explicit endorsement of prejudiced beliefs about the target group in general, such that evaluative associations and explicit prejudice were correlated only for participants with a low motivation to control prejudiced reactions, but not for those high in motivation to control. Taken together, these results suggest that perceivers sometimes infer different attributes from one and same behaviour, depending on the actor's category membership. Most importantly, this seems to be true even when perceivers do not explicitly endorse prejudiced beliefs about the target group. In other words, our impressions of a stereotyped target can still be prejudiced, even when we are highly motivated to react in an unprejudiced manner.

ACKNOWLEDGEMENTS

This research was supported by a grant from the Deutsche Forschungsgemeinschaft to Rainer Banse (Ba 1603/1-1). Portions of this article were presented at the 3rd European Social Cognition Network Meeting (Houffalize, Belgium), which was supported by a grant from the European Union. We are grateful to Galen Bodenhausen, Ap Dijksterhuis, Russel Fazio, Kurt Hugenberg, Alan Lambert, Fritz Strack, and three anonymous reviewers for helpful comments on an earlier version of this draft; Iain Glen for stylistic corrections; and Iris Petzold for her help in collecting the data.

REFERENCES

Banse, R., & Gawronski, B. (2003). Die Skala Motivation zu vorurteilsfreiem Verhalten: Psychometrische Eigenschaften und Validität [The scale motivation to act without prejudice: psychometric properties and validity]. *Diagnostica, 49*, 4–13.

Banse, R., Seise, J., & Zerbes, N. (2001). Implicit attitudes towards homosexuality: Reliability, validity, and controllability of the IAT. *Zeitschrift für Experimentelle Psychologie, 48*, 145–160.

Bargh, J. A., Gollwitzer, P. M., Lee-Chai, A., Barndollar, K., & Trötschel, R. (2001). The automated will: Nonconscious activation and pursuit of behavioural goals. *Journal of Personality and Social Psychology, 81*, 1014–1027.

Beringer, J. (1994). *Experimental Run Time System 3.18.* Frankfurt, Germany: BeriSoft Cooperation.

Blair, I. V. (2001). Implicit stereotypes and prejudice. In G. B. Moskowitz (Ed.), *Cognitive social psychology* (pp. 359–374). Mahwah, NJ: Erlbaum.

Blair, I., Ma, J., & Lenton, A. (2001). Imagining stereotypes away: The moderation of implicit stereotypes through mental imagery. *Journal of Personality and Social Psychology, 81*, 828–841.

Brewer, M. B. (1988). A dual process model of impression formation. In T. K. Srull, & R. S. Wyer (Eds.), *Advances in social cognition* (Vol. 1, pp. 1–36). Hillsdale, NJ: Erlbaum.

Brewer, M. B., & Brown, R. (1998). Intergroup relations. In D. T. Gilbert, S. T. Fiske, & G. Lindsey (Eds.), *Handbook of social psychology* (4th ed., Vol. II, pp. 554–594). New York: McGraw-Hill.

Bruner, J. S. (1957). On perceptual readiness. *Psychological Review, 64*, 123–152.
Darley, J. M., & Gross, P. H. (1983). A hypothesis-confirming bias in labeling effects. *Journal of Personality and Social Psychology, 44*, 20–33.
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). Stereotype and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56*, 5–18.
Dovidio, J. F., Kawakami, K., & Beach, K. R. (2001). Implicit and explicit attitudes: Examination of the relationship between measures of intergroup bias. In R. Brown, & S. L. Gaertner (Eds.), *Blackwell handbook of social psychology: Intergroup processes* (pp. 175–197). Malden, MA: Blackwell.
Duncan, B. L. (1976). Differential perception and attribution of intergroup violence: Testing the lower limits of stereotyping of Blacks. *Journal of Personality and Social Psychology, 34*, 590–598.
Dunning, D., & Sherman, D. A. (1997). Stereotypes and tacit inference. *Journal of Personality and Social Psychology, 73*, 459–471.
Dunton, B. C., & Fazio, R. H. (1997). An individual difference measure of motivation to control prejudiced reactions. *Personality and Social Psychology Bulletin, 23*, 316–326.
Fazio, R. H. (1990). Multiple processes by which attitudes guide behaviour: The MODE model as an integrative framework. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 75–109). San Diego, CA: Academic Press.
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* (pp. 247–282). Mahwah, NJ: Erlbaum.
Fazio, R. H., & Williams, C. J. (1986). Attitude accessibility as a moderator of the attitude-perception and attitude-behaviour relations: An investigation of the 1984 presidential election. *Journal of Personality and Social Psychology, 51*, 505–514.
Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology, 69*, 1013–1027.
Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 1–74). San Diego, CA: Academic Press.
Gawronski, B. (2002). What does the Implicit Association Test measure? A test of the convergent and discriminant validity of prejudice related IATs. *Experimental Psychology, 49*, 171–180.
Gawronski, B., Alshut, E., Grafe, J., Nespethal, J., Ruhmland, A., & Schulz, L. (2002). Prozesse der Urteilsbildung über bekannte und unbekannte Personen: Wie der erste Eindruck die Verarbeitung neuer Informationen beeinflusst [Processes of judging known and unknown persons: How the first impression influences the processing of new information]. *Zeitschrift für Sozialpsychologie, 33*, 25–34.
Gawronski, B., Ehrenberg, K., Banse, R., Zukova, J., & Klauer, K. C. (2003). It's in the mind of the beholder: The impact of stereotypic associations on category-based and individuating impression formation. *Journal of Experimental Social Psychology, 39*, 16–30.
Greenwald, A. G., & Nosek, B. A. (2001). Health of the Implicit Association Test at age 3. *Zeitschrift für Experimentelle Psychologie, 48*, 85–93.
Greenwald, A. G., McGhee, D. E., & Schwartz, J. K. L. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology, 74*, 1464–1480.
Higgins, E. T., Rholes, W. S., & Jones, C. R. (1977). Category accessibility and impression formation. *Journal of Experimental Social Psychology, 13*, 141–154.
Houston, D. A., & Fazio, R. H. (1989). Biased processing as a function of attitude accessibility: Making objective judgments subjectively. *Social Cognition, 7*, 51–66.
Jackson, L. A., Sullivan, L. A., & Hodge, C. N. (1993). Stereotype effects on attributions, predictions, and evaluations: No two social judgments are quite alike. *Journal of Personality and Social Psychology, 65*, 69–84.
Kawakami, K., Dovidio, J. F., Moll, J., Hermsen, S., & Russin, A. (2000). Just say no (to stereotyping): Effects of training in the negation of stereotypic associations on stereotype activation. *Journal of Personality and Social Psychology, 78*, 871–888.
Klauer, K. C., & Wegener, I. (1998). Unraveling social categorization in the 'Who said what?' paradigm. *Journal of Personality and Social Psychology, 75*, 1155–1178.

Kunda, Z., & Sherman-Williams, B. (1993). Stereotypes and the construal of individuating information. *Personality and Social Psychology Bulletin, 19*, 90–99.

Kunda, Z., & Thagard, P. (1996). Forming impressions from stereotypes, traits, and behaviours: A parallel constraint satisfaction theory. *Psychological Review, 103*, 284–308.

Lepore, L., & Brown, R. (1997). Category and stereotype activation: Is prejudice inevitable? *Journal of Personality and Social Psychology, 72*, 275–287.

Moskowitz, G. B., Gollwitzer, P. M., Wasel, W., & Schaal, B. (1999). Preconscious control of stereotype activation through chronic egalitarian goals. *Journal of Personality and Social Psychology, 77*, 167–184.

Moskowitz, G. B., Salomon, A. R., & Taylor, C. M. (2000). Preconsciously controlling stereotyping: Implicitly activated egalitarian goals prevent the activation of stereotypes. *Social Cognition, 18*, 151–177.

Newman, L. S., & Uleman, J. S. (1990). Assimilation and contrast effects in spontaneous trait inference. *Personality and Social Psychology Bulletin, 16*, 224–240.

Otten, S., & Moskowitz, G. B. (2000). Evidence for implicit evaluative in-group bias: Affect-biased spontaneous trait inference in a minimal group paradigm. *Journal of Experimental Social Psychology, 36*, 77–89.

Pettigrew, T. F., & Meertens, R. W. (1995). Subtle and blatant prejudice in western Europe. *European Journal of Social Psychology, 25*, 57–75.

Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology, 75*, 811–832.

Rudman, L. A., Ashmore, R. D., & Gary, M. L. (2001). 'Unlearning' automatic biases: The malleability of implicit prejudice and stereotypes. *Journal of Personality and Social Psychology, 81*, 856–868.

Sagar, H. A., & Schofield, J. W. (1980). Racial and behavioural cues in black and white children's perceptions of ambiguously aggressive acts. *Journal of Personality and Social Psychology, 39*, 590–598.

Schuette, R. A., & Fazio, R. H. (1995). Attitude accessibility and motivation as determinants of biased processing: A test of the MODE model. *Personality and Social Psychology Bulletin, 21*, 704–710.

Schwibbe, M., Räder, K., Schwibbe, G., Borchardt, M., & Geiken-Pophanken, G. (1994). Zum emotionalen Gehalt von Substantiven, Adjektiven und Verben [The emotional meaning of nouns, adjectives, and verbs]. In W. Hager, & M. Hasselhorn (Eds.), *Handbuch deutschsprachiger Wortnormen* (S. 272–284). Göttingen: Hogrefe.

Srull, T. K., & Wyer, R. S. (1979). The role of category accessibility in the interpretation of information about persons: Some determinants and implications. *Journal of Personality and Social Psychology, 37*, 1660–1672.

Srull, T. K., & Wyer, R. S. (1980). Category accessibility and social perception: Some implications for the study of person memory and interpersonal judgments. *Journal of Personality and Social Psychology, 38*, 841–856.

Steffens, M. C., & Plewe, I. (2001). Item's cross-category associations as a confounding factor in the Implicit Association Test. *Zeitschrift für Experimentelle Psychologie, 48*, 123–134.

Strack, F., & Hannover, B. (1996). Awareness of the influence as a precondition for implementing correctional goals. In P. M. Gollwitzer, & J. A. Bargh (Eds.), *The psychology of action: Linking cognition and motivation to behaviour* (pp. 579–596). New York: Guilford Press.

Taylor, S. E., Fiske, S. T., Etcoff, N. L., & Ruderman, A. J. (1978). Categorical and contextual bases of person memory and stereotyping. *Journal of Personality and Social Psychology, 36*, 778–793.

Trope, Y. (1986). Identification and inferential processes in dispositional attribution. *Psychological Review, 93*, 239–257.

Trope, Y., & Alfieri, T. (1997). Effortfulness and flexibility of dispositional judgment processes. *Journal of Personality and Social Psychology, 73*, 662–674.

Trope, Y., Cohen, O., & Maoz, Y. (1988). The perceptual and inferential effects of situational inducements on dispositional attribution. *Journal of Personality and Social Psychology, 55*, 165–177.

Trope, Y., Cohen, O., & Alfieri, T. (1991). Behaviour identification as a mediator of dispositional inference. *Journal of Personality and Social Psychology, 61*, 873–883.

Uleman, J. S., Newman, L. S., & Moskowitz, G. B. (1996). People as flexible interpreters: Evidence and issues from spontaneous trait inference. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 28, pp. 211–279). San Diego, CA: Academic Press.

University of Stirling Psychology Department (2000). *Psychological Image Collection at Stirling*. http://pics.psych.stir.ac.uk/index.html.

Wegener, D. T., & Petty, R. E. (1997). The flexible correction model: The role of naive theories of bias in bias correction. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 29, pp. 141–208). San Diego, CA: Academic Press.

Wittenbrink, B., Judd, C. M., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationships with questionnaire measures. *Journal of Personality and Social Psychology, 72*, 262–274.

Wittenbrink, B., Judd, C. M., & Park, B. (2001). Spontaneous prejudice in context: Variability in automatically activated attitudes. *Journal of Personality and Social Psychology, 81*, 815–827.
Zick, A. (1997). *Vorurteile und Rassismus. Eine sozialpsychologische Analyse [Prejudice and Racism: A social psychological analysis].* Münster: Waxmann.

APPENDIX

Scenarios Used To Assess Dispositional Inference

Situation 1: One evening, A is walking along a small, quiet street. There is no one on the street except an old man walking a few metres in front of A. Close to a street lamp, the old man's wallet falls out of his pocket without him noticing it. How likely do you think is it that A grabs the wallet and does not return it to the old man?
Situation 2: A is at the train station and waits for his train to come. On arrival of the train, a big crowd emerges and a man hits A strongly with his suitcase. How likely do you think it is that A will get enraged in response to this event?
Situation 3: A uses the public transportation very infrequently. One afternoon he is surprised when some officers ask for his ticket. How likely do you think is it that A does not have a valid ticket?
Situation 4: A is sitting in a cafe and is asked by a student to take part in a psychological study. In this study A is asked to respond to some questions about his romantic relationship. Because he has nothing else to do, A agrees to take part in this study. How likely do you think it is that A does not give true answers to the questions?
Situation 5: After a very stressful work day A is driving home in his car, when he got stuck in a traffic jam. After a few minutes the cars start moving again and another car next to him changes lanes right in front him, such that A nearly hits the other car. How likely do you think it is that A will react very aggressive to the other driver?
Situation 6: One cold winter evening A is on his way back home from a party. It's about midnight and he is a bit drunk. On the way to his car, he considers whether to walk two kilometres to his home, or to drive with his car. How likely do you think it is that A drives even though he is drunk?
Situation 7: A few days ago A was surprised to meet an old friend from high school. They immediately arranged a meeting at his friend's place. Now, A is at his friend's place when he is asked by his friend whether he would like to smoke some marihuana. How likely do you think it is that A accepts this offer?
Situation 8: At the train station the escalators don't work. Hence, A starts up the stairs when he sees an old woman with a heavy suitcase walking up in front of him. How likely do you think it is that A passes the old woman without asking her to help her with the suitcase?
Situation 9: A is waiting for the subway on his way back from the supermarket. On arrival of the train, the handle of one of his bags rips and a milk bottle falls out of the bag and smashes on the ground. How likely do you think it is that A yells at another person next to him even though this person is not responsible for the ripped bag?

