On the nature and expression of ethnic prejudice
as seen in judgments of pictorial stimuli

Bert Westerlundh, Sven Birger Hansson, Fredrik Björklund, and Martin Bäckström

Department of Psychology, Lund University

In Study 1 (N = 120), pictures of violent behavior taking place within or between ethnic groups were rated under different instructions. A strong cross-over effect was found: interracial violence was reported as relatively more threatening under the instruction that the study concerned anxiety, whereas violence among Whites was reported as more threatening under the instruction that the study concerned stereotyping. In a separate task participants made trait ratings based on faces of members of different ethnic groups. Black men (but not Black women) received very positive ratings. Study 2 (N = 58) replicated this finding, showing it to be unrelated to both self-reported motivation to appear unprejudiced and to prevalent naive theories on how different ethnic groups are evaluated (theory-based correction). Results are discussed in terms of correction mechanisms and self-presentation concerns.

Key words: prejudice, self-presentation, overcorrection

The present paper concerns ethnic prejudice as revealed in responses to pictorial stimuli, and how expression of prejudice is influenced by self-presentation concerns. Self-presentation concerns have become important in research on the relationship between implicit and explicit prejudice (Dunton & Fazio, 1997; Nosek, 2002). In a typical dual-process account (see the volume edited by Chaiken & Trope, 1999), implicit prejudice is activated on an automatic (unconscious, non-volitional, parallel) level of processing by a social stimulus. The product of this automatic process can be inhibited and remain unconscious. We thus arrive at Devine's (1989) formula that stereotype activation is uncontrollable, while stereotype application might be controllable. Persons who report having a strong motivation to inhibit prejudicial responses

will do so if they have the opportunity, and produce something else than the automatic negative evaluation as a conscious report. Dunton and Fazio (1997) presented an introspective self-report measure of this motivation to appear unprejudiced, and Fazio and Dunton (1997) found support for the automatic – controlled response model in a clever experiment. Several experimental findings in this area seem not to be the product of data-driven, individuating processing, but rather to be crude overcompensations of negative judgments (Fazio, Jackson, Dunton & Williams, 1995; Guimond, Dambrun, Michinov & Duarte, 2003, and von Hippel, Silver & Lynch, 2000; for a review see Bodenhausen & Moreno, 2000). There is also growing experimental evidence for unconscious inhibition, that implicit prejudiced attitudes can be countered by unconscious intentions (Araya, Akrami, Ekehammar & Hedlund, 2002; Bargh, Gollwitzer, Lee-Chai, Barndollar & Trötschel, 2001). This may be linked to the fact that the motive to present a positive view of one's self, to others and to oneself, is strong and influences many areas of life, a finding which is extremely well substantiated empirically (Greenwald & Banaji, 1995). A number of biases are involved when the self is concerned, and as regards suppression of unacceptable contents, researchers in this area have long differentiated between conscious impression management and unconscious self deception (Millham & Kellogg, 1980; Paulhus & Reid, 1991).

There is a literature on correction processes in social judgments, with models varying in assumptions concerning the cognitive mechanisms and the level of conscious control involved. Wegner's (1994) model of ironic control includes an unconscious low resource-demanding monitor. If the intention is to suppress, the monitor will search for items which are the objects of suppression, and warn the operating system of their presence. The operating system is the conscious intention to suppress and the attentional resources allotted to it. In suppression, it is essentially a distraction device. It directs attention to anything that does not belong to the forbidden sphere. Available resources are critical in this account. If they become low the operating system may fail, and the forbidden items on which the monitor focuses become conscious. This model of impression management (and its failure) has its roots in experiments where subjects are directed to suppress thoughts, which is not the case in the present study.

In contrast, Wegener and Petty's (1995) model of flexible correction stems from a tradition of priming experiments. The central dimension is one of target assimilation to versus contrast to the prime, and the model describes how individuals engage in attempts to correct for the influence of the prime. Depending on a person's naive theory of how the context affects judgments, corrections can go either way. These theories can be stated as propositions:

B is better than A; if I evaluate A in a context involving B, I will give A a very low rating. Therefore, in that situation I must correct my evaluation of A upwards. Persons can be asked to describe their naive theories. This is again a theory of impression management, but one which describes active attempts at compensation.

Finally, the social defense theory described by Baumeister, Dale, and Sommer (1998) is a development of the Freudian theory of defense mechanisms. The motive for defensive activity here is protection of self-esteem. The theory applies to normal people and to moderate rather than extreme forms of defense. In their review of social psychological research Baumeister et al find reaction formation (overcompensation), isolation, and denial to be well demonstrated. These mechanisms are resource-demanding but often act outside awareness. As opposed to the flexible correction model this theory recognizes that we are often unconscious self-deceivers.

On the whole, these models are actually not opposed to each other but complementary. As regards unconscious self-deception, the regulation of self-presentation is a major concern from early years on (Banerjee, 2002), and thus an activity where automatization should take place. It is an example of what Forgas (2000) calls the motivated processing strategy, limited and directed by a specific motivational objective. It may well be one of the major instances of the automated will (Bargh et al., 2001). The present study investigates the importance of such automatic correction processes in evaluations of outgroups. The stereotypes concerning outgroups may present them as dangerous and aggressive. The seminal study by Devine (1989) on the automatic stereotype of the African American was concerned with exactly this theme. The appraisal of danger involves an estimation of one's own power relative to the object. Those who feel low in power should regard outgroup members as more dangerous. In this regard, women would be expected to show more fear-based racism than men. This is not what is found in questionnaire studies in different societies, where men consistently display more racial prejudice than women (Ekehammar, Akrami & Araya, 2003).

Our first study is a full-scale trial run, designed as a factorial field experiment (with two factors, sex of subjects and instruction). There are two dependent measures. In one, participants rate violence-related pictures on dimensions of dangerousness. Half of the pictures depict violence between Whites, half of them interracial violence. The saliency of the ethnic content is varied through instructions. For half the participants, the test is presented as a measure of anxiety. The other half get the instruction that the study is about ethnic stereotypes. Of particular interest is a possible cross-over effect: that interracial violence will be reported as more threatening than ingroup violence when the racial content is hidden, but

as less threatening when the racial content is stressed. This will give a measure of the strength of the social desirability pressures involved in questions concerning racism. Nosek (2002) has pointed out that too little experimental work has been done in this area. Further, it is possible to compare the reports of the sexes to these violence-related pictures. As stated, females are expected to be more prone to fear-related racism.

The second measure also uses pictorial stimuli, in this case exemplars of three racial groups, Blacks, Whites and Asians. The stimuli vary along two further dimensions: sex (male and female) and looks (for each race and sex category there is one attractive and one not so attractive exemplar). In this case, half of the participants are told that the instrument measures general attitudes based on first impressions of peoples' looks, the other half that it is a measure of ethnic stereotypes. Since race is a highly accessible category subject to self-presentation concerns, it is not too likely that these instructions will have an effect. But correction processes, perhaps resulting in overcompensation in responses to Blacks (as described above, Dunton & Fazio, 1997; Wegener & Petty, 1995) should be at work in both instruction groups. No strong prejudice toward people with Asian background is expected. Most of those seen by our participants are likely to be adopted children brought up in Sweden.

Study 1

Method

Participants

The participants were one hundred and twenty native Swedish high-school pupils, 60 from each of two schools in southern Sweden. In each school group there were 30 males and 30 females, aged 16 to 17. They took part in the investigation without compensation.

Materials

The faces test. This instrument consisted of twelve black-and-white photographs of faces varying in sex (Male – Female), looks (Good – Less good), and ethnicity (Black, Asian and White). They were taken from published sources. Using the capitalized initials, the photographs were presented to all participants in the following order: FGB, MLA, FLW, MGW, FLA, MGB, FGA, MLB, FGW, FLB, MGA, and MLW. There were three photographs to one page in the booklet, each approximately 8 (h) x 7.5 (w) cm. All photographs were rated by circling a number from 1 to 7 on ten dimensions. Ten pairs of adjectives were used to define the dimensions (Appendix A). For five items, 1 was the most

positive score, for five it was the most negative. In scoring, the positive were reversed so that higher scores always indicated a more negative evaluation, and all ten dimensions then were aggregated into an index.

The violence test. This instrument consisted of six pictures from old comic strips. All of them included threat or overt violence. The second, third and sixth presented interracial violence, the first, fourth and fifth violence between Whites. The response format was the same as in the faces test with two pictures approximately 7 (h) cm to a page. The adjective pairs used to evaluate the pictures are presented in Appendix A. The scoring was performed so that higher scores indicated a more negative view of a picture (higher perceived threat). As for the faces test the rating dimensions were aggregated into an index.

Procedure

The instruments were collected in a booklet. The first part contained the instructions. Thirty males and 30 females at each school got the instruction that the study was about "stereotypes, the ideas we have of different groups of people" (the stereotype instruction). The other half got the instructions that the study concerned "halo effects, very general attitudes, and if scenes are interpreted as threatening. I believe this influences our social attitudes. To study if persons have very general attitudes one can see if their evaluation of people is influenced by how these people look. To study if persons interpret scenes as threatening, one can ask them to rate pictures." (the halo/anxiety instruction). Then followed the tests, always in the order of the faces test followed by the violence test.

Following some additional measures that will not be reported on here, the last page of the booklet told the participants that attentiveness could have influenced their way of responding. Therefore, they were asked to state in their own words what they thought was the purpose of the study. This check was included to see if any pupil with the halo/anxiety instruction would write that this was an investigation focused on discrimination, stereotypes or racism. There was no such pupil. On the declarative level, none reported the instruments to specifically measure prejudice. The instruments were administered by one of the pupils' ordinary teachers during a class hour.

Results

The violence test

Descriptive results for the violence test are found in Table 1.

Table 1

Mean threat ratings in the violence test, across sex of participant

Violence	Male (n = 60)		Female (n = 60)		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*
White	5.51	1.00	5.69	.96	5.60	.98
Interracial	5.61	1.00	6.15	1.03	5.88	1.04
White minus Interracial	-.10	1.17	-.45	1.11	-.28	1.15

Following the recommendations of Guimond, Dambrun, Michinov and Duarte (2003), a difference score of White minus Interracial violence was calculated. This was negative for the total group; $t(119) = -2.64$, $p < .01$, $\eta^2 = .06$. In general, Interracial violence was seen as more threatening than White violence. To test the main hypothesis a two-factor ANOVA was conducted with Sex (female vs. male) and the Instruction condition (anxiety vs. stereotype) as between group factors and the difference score as the dependent variable. The main effect of Sex was marginally significant, $F(1, 116) = 3.88$, $p = .051$, $\eta^2 = .03$, females making higher threat ratings than males. The instructions created a powerful cross-over effect (Figure 1).
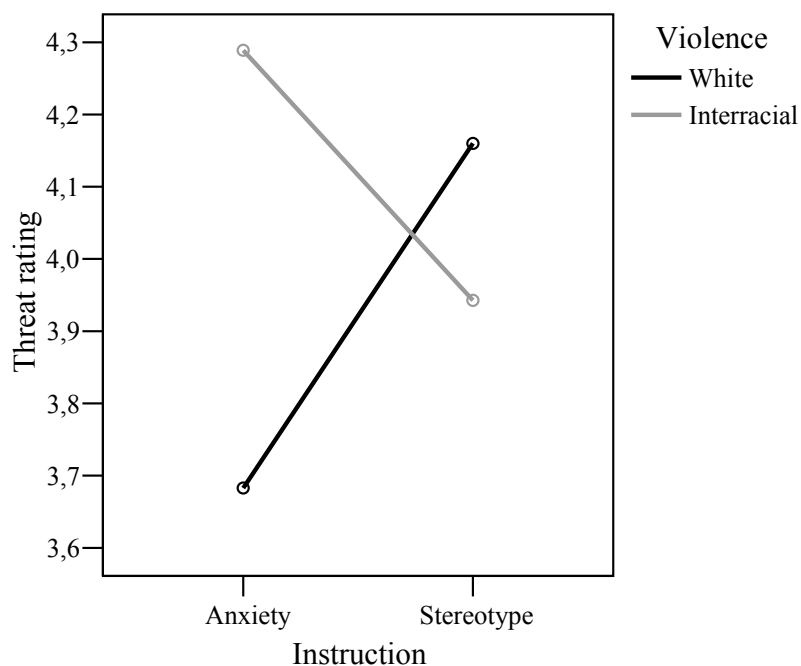
Figure 1. Threat ratings as an effect of violence group and experimental instruction.

Interracial violence was reported as relatively more threatening under the halo/anxiety instructions, while White violence was reported as more threatening under the stereotype instructions, $F(1, 116) = 42.86, p < .001, \eta^2 = .27$. Both the ratings of Interracial violence and the ratings of White violence changed significantly between conditions; for Interracial, $t(118) = -2.67, p < .01, \eta^2 = .06$; for White, $t(118) = 4.05, p < .001, \eta^2 = .12$). The interaction between sex and instruction condition was non-significant ($F < 1$). Females rated Interracial violence as more threatening than males in both conditions.

The hypotheses for this experiment concerned first the power of social desirability pressures when the topic of prejudice is made salient in a social judgment situation as compared to when it is not. Our expectation that this power would be great was supported by the results of the experiment (partial $\eta^2 = .27$, corresponding to a Cohen's d of 1.19). It turned out that compensation efforts were not limited to the evaluation of the outgroup but were even stronger in the opposite direction for the ingroup (Whites). Clearly evaluation of both outgroup and ingroup can be affected by compensation efforts. The second hypothesis referred to the stereotype of outgroups as aggressive. Females gave marginally more negative reports to such stimuli, as compared to males.

The faces test

Results for the two instructions groups are found in the two first columns of Table 2.

Table 2.

Mean ratings for faces differing in sex and race, across experimental condition

Faces	Halo		Stereotype		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*
Black women	3.72	.69	3.59	.70	3.65	.70
Black men	5.08	.76	4.96	.71	5.02	.74
White women	4.42	.64	4.52	.66	4.47	.65
White men	3.66	.61	3.70	.52	3.68	.56
Asian women	4.14	.58	3.92	.63	4.02	.61
Asian men	3.94	.52	3.92	.60	3.93	.56

Note: High values indicate a positive rating.

ANOVAs on ratings of the pictures of the different sexes and races revealed no significant results concerning sex of participant. There was only one significant effect of instruction condition, the lower value for Asian women in the stereotype as compared to the halo/anxiety instruction condition, $F(1, 112) = 3.93$, $p < .05$, $\eta^2 = .03$.

However, the ratings of the various types of faces differed. A similar contrast as in the picture test was created, based on the ratings of all the White faces against all the Asian and Black faces. This contrast depicts the degree of preference for White as compared to Asian and Black faces. It could be expected that the White faces would be preferred to a lesser degree under the stereotype instruction, but the opposite was found. Participants under the stereotype instruction rated the White faces more positively than the Black/Asian faces, $t(118) = 2.29$, $p < .05$, $\eta^2 = .04$.
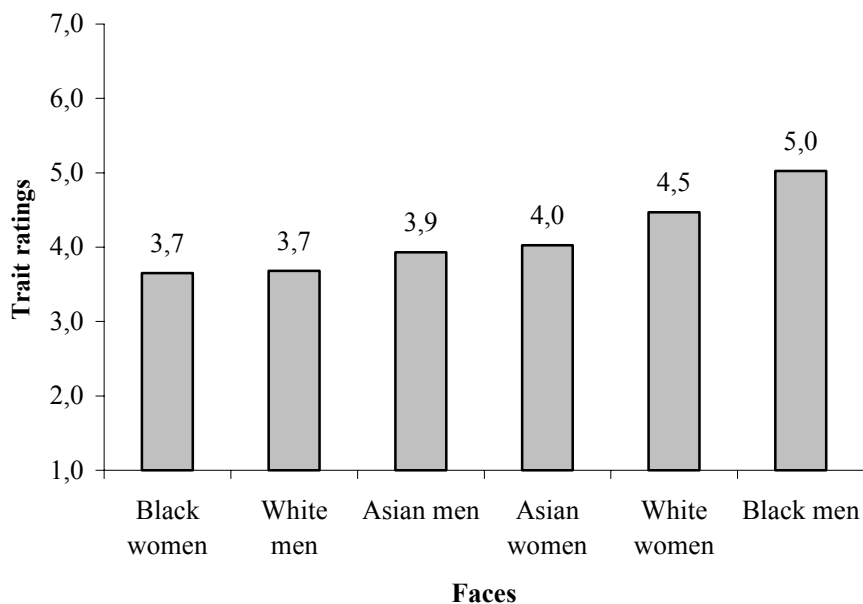

Figure 2. Mean positive trait ratings in the faces test (Study 1).

Figure 2 displays the ratings of all six types of faces. Many ratings were significantly different from each other, as tested by pairwise t-tests. Generally, differences larger or equal to scores of .2 were significant at $p < .05$. This makes for example the ratings of White men significantly more negative than the ratings of Asian men. Black men clearly received the highest ratings, followed by White women, and Asian women. Black women together with White men received significantly lower ratings than all other categories of faces. The ratings

of the Asian faces fell in between the other faces. The conspicuous finding, of course, is the positive ratings of the Black men.

Discussion

The results on the faces test are in accordance with Wegener and Petty's (1995) theory of attitude correction. The stereotype of the male gives potency and activity to this sex, which could make out-group members of this sex more threatening and more disliked. This could be the basis for naive theories underlying the overcompensation effect. Such an account leaves out other possibilities, however, and requires further study.

In two samples, one of which was made aware that the study concerned stereotypes, we have found a pattern of ratings of Blacks, the physically most different outgroup, which seems to reflect a simple heuristic. Socially potent (male) members of the outgroup received overcompensated ratings, while no overcompensation was seen for the socially less potent (female) members. The values for Black women were not outside the range of ratings of other groups – they were almost equal to those for White men. If Blacks had shown the same man – woman difference as other groups, Black men would have been the most negatively rated of all groups. Instead, the opposite was found.

Study 2

Research on the expression of racial prejudice in the United States during the last half of the twentieth century has shown a strong increase in Whites' reported support for racial equality. Blatantly racist expressions are socially condemned. But Whites do not show a corresponding increase in support of concrete policies intended to bring about greater racial equality (Schuman, Steeh & Bobo, 1985). Racism has not disappeared; it has taken on new, more disguised forms. A distinction has been made between "classical", blatant, undisguised racism and "modern", subtle and indirect racism.

Classical racism was defined by McConahay (1986) as a strong antiegalitarian negative evaluation of an outgroup, combined with a powerful motive to keep members of this group from status/power advancement. Questionnaire items that measure classical racism all stress the inherent inferiority of the outgroup. Modern racism is a descriptive term for subtle forms of racial prejudice, and a number of influential accounts have pointed to different aspects of this phenomenon.

Dovidio and Gaertner's (1998) account of aversive racism stresses inhibition of unconscious racism and its consequences. It concerns persons who have a strong negative

evaluation of the outgroup at an implicit level, and at the same time consciously uphold modern ideas of justice and egalitarianism. Aversive racists deny prejudice and uphold this denial by different forms of suppression. The conflict leads to avoidance of race-related topics, and of contacts with members of the outgroup. If aversive racists are forced to engage in such contacts, their behavior is inhibited and avoidant, which is felt negatively by the outgroup persons involved. Avoidant racists may well engage in discriminatory behavior, if the negative evaluation of the outgroup members can be attributed to factors other than race. This could be seen as an elaboration of a position where modern racism is basically of the same nature as classical racism, but diluted and inhibited (compare McConahay, 1986). In a Scandinavian context, Akrami, Ekehammar and Araya (2000) developed scales of classical and modern racism. A response suppression dimension seems to be the central one in their thinking, in line with that presented above. They found a strong positive correlation between the scales, but concluded from confirmatory factor analysis that the distinction between classical and modern racism is viable also in Scandinavia.

In opposition to this is the concept of symbolic racism (Kinder & Sears, 1981). The new type of racism is seen as qualitatively different from the old one. It is not based on anti-egalitarianism, but rather on negative affect toward the outgroup in combination with individualism and moral conservatism (the Protestant work ethic, Kinder 1986; Sears, 1988). Outgroup members are seen as parasites who live on the social welfare system, instead of taking responsibility for their lives and work for a living. They are thus condemned as immoral, and egalitarian measures to better their position are seen as unjust. This theory has been criticized on different grounds. The idea of a connection between individualism and modern racism has not been empirically substantiated (Sniderman & Tetlock, 1986a; 1986b). The values involved ("American individualism") are limited to a specific culture (Sidanius & Pratto, 1999). Yet, the theory seems to capture the flavor of modern racism. Sears (1988) distinguished three components of it: denial of continued discrimination, antagonism toward minority group demands, and resentment about special favors for minority groups. In line with what has been presented above, questionnaire items on modern racism deny that minority groups have special problems, deny that any form of action needs to be taken on their behalf, and state that anyway, one hears too much about these groups and their so-called problems.

Study 2 is an attempt to replicate the correction-related finding from Study 1. The validity of the faces test is studied by relating it to measures of modern and classical racism.

But first we present two preliminary studies on possible mechanisms behind the correction effect.

Method

Preliminary study 1: Is the reason that Black men received more positive trait ratings in the faces test that they are perceived as more attractive? The faces used in Study 1 varied on race, sex and looks. Faces in each race/sex combination varied in looks (attractive versus less attractive), while looking normal and reasonably ordinary for their race group. Sex and race were in fact the powerful variables in the study, while the difference in looks apparently played a minor part. Pertinent information regarding this is presented in Figure 3.
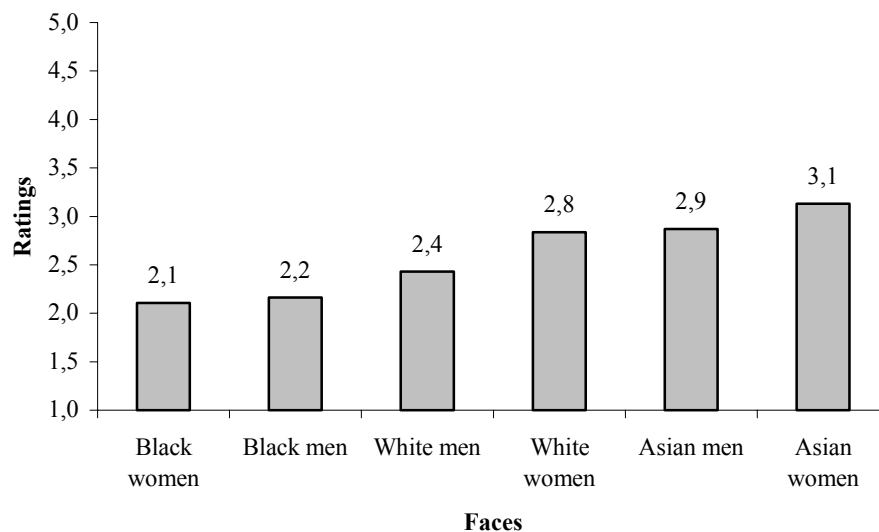
Figure 3. Mean ratings of the faces on the attractiveness/well-groomedness index.

Participants are thirty students in their early twenties. They rated the faces on attractiveness and well-groomedness on scales ranging from 1 to 5. Figure 3 depicts the index composed of the means of the two ratings. Higher values indicate a more positive rating. The figure shows that Black men were *not* favored on these ratings, as opposed to their trait ratings in Study 1. The difference between Black men and White men, for example, was not significant ($t < 1$). Ratings of Black women and Asian women showed the largest difference.

Preliminary study 2: Were the positive ratings of Black men in Study 1 due to participants' naive theories, as Wegener and Petty's (1995) model proposes, or were they rather a crude response to negative affect felt in the situation? To study this, a group of

students was asked to state their naive theories, that is, their beliefs concerning our "real" and "basic" attitudes to different groups of people, disregarding social desirability demands. There were 89 participants, 60 women and 29 men, all students with a mean age in their early twenties. 83.1 per cent believed that Whites are rated more positively than Blacks. 25.8 per cent believed that Black men and Black women are rated equal. 21.3 per cent thought that Black women are rated more positively than Black men, while 52.8 per cent thought that Black men are rated more positively than Black women. Based on the naive theories, there are grounds for applying compensation in order not to appear prejudiced. But the need for compensation would be stronger for ratings of Black women than for ratings of Black men. The results in the faces test are thus not in accordance with the theory-based correction view, but rather point to situational inhibition of negative affect. This is in turn related to limitations in the control of negative evaluations of the outgroup, a topic discussed by Bargh (1999). In order to investigate further whether socially desirable evaluation of the faces is related to conscious control processes, the Dunton & Fazio (1997) measure of individual differences in the motivation to control prejudice was included in Study 2. This scale has to do with introspectively available concerns to avoid dispute and to appear unprejudiced. Is this the type of control which will be found in the ratings of pictures?

Participants

Fifty-eight persons, 32 women and 26 men, volunteered to participate in the study without compensation. The age range was 18 to 51, with a mean of 23.5.

Materials

The faces were presented in a different order than in Study 1. With the notation used there, the faces were presented in the order FGB, MLA, FGW, FGA, MLB, FLW, FLB, MGA, MLW, MGW, FLA, and MGB. The response format and scoring were the same as in Study 1. Further, the Scandinavian versions of the modern and classical racial prejudice scales of Akrami et al. (2000) were used. The seventeen items of these scales were mixed with fifteen filler items that concerned social issues. Ratings were made on a Likert-scale ranging from 1 *(disagree completely)* to 5 *(agree completely)*. Scoring followed the published instructions, so that higher values always corresponded to higher racial prejudice. Finally, a Swedish version of Dunton and Fazio's (1997) motivation to control prejudice scale was included. Response format (-3, disagree completely to +3, agree completely) and scoring

followed the published instructions, and the reliability was found to be satisfactory (Cronbach's *alpha* = .79).

Procedure

Testing took place in a laboratory, with some other instruments included which are not reported upon here. The number of participants tested at a time varied, the maximum was seven. Participants were seated in individual booths and told that "the tests measure your social attitudes to people of different ethnic origins". The order of the faces test and the block of racial prejudice tests was rotated between subjects. Dunton and Fazio's scale was always presented last.

Results

This study was an extended replication of the presumed correction effect from the faces test in Study 1. Based on the previous results we predicted that Black men would be rated higher than all other groups. Figure 4 displays the results and it is readily seen that Black men were regarded as having more positive traits than all other groups (all *p*s < .05). As in Study 1 Black women were rated much lower than Black men.
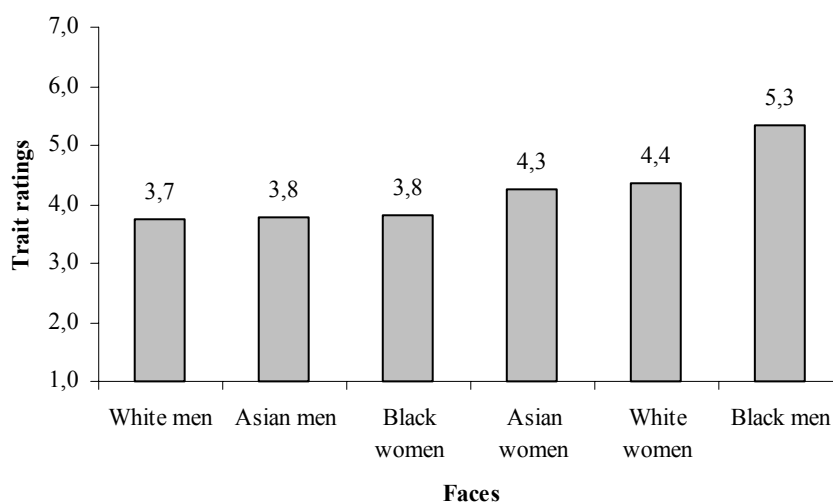
Figure 4. Mean positive trait ratings in the faces test (Study 2).

Dunton and Fazio's motivation to control prejudiced reactions scale. Motivation to control should be related to the ratings of Black and Asian faces, but not neccessarily to the White faces. Results are found in Table 3.

Table 3.

Correlations between face ratings, the motivation to control prejudice, and racism

	Black women	Black men	White women	White men	Asian women	Asian men	Prejud. control	Modern racism
Black men	.30*							
White women	-.26	.19						
White men	.05	.18	.22					
Asian women	.54**	.36**	-.14	.00				
Asian men	.27*	.16	.02	.22	.15			
Prejudice control	.12	.21	-.02	.00	.24	.08		
Modern racism	-.46**	-.40**	.24	.13	-.25	-.03	-.36**	
Classical racism	-.50**	-.35**	.21	.06	-.25	-.18	-.26*	.79**

Note: * = sign. at *p* < .05; ** = sign. at *p* < .01.

The correlations were generally low and non-significant. Motivation to control seems not to have influenced the face ratings. An alternative analysis, with somewhat more power, would be to look at participants who rated the Black men very high but all other foreign faces low, to see if their ratings were influenced by the motivation to control prejudice. A variable was created based on this hypothesis and was tested against the Dunton and Fazio scale. The correlation was close to zero. All possible variables where Black men are contrasted to other groups were constructed and related to the measure of motivation to control prejudice, but correlations were weak. In other words, motivation to control could not explain the high ratings of Black men.

Modern and classical racial prejudice. The correlation pattern indicated that Race was an important factor behind the ratings in the faces test. The modern and classical racial prejudice scales correlated strongly and positively with each other (Table 3). Despite the fact that in the faces test Black men were rated high and Black women rather low, both had significant negative correlations with the prejudice scales. Further, the ratings of Black men and Black women were significantly positively related. In addition, there were positive correlations between ratings of Black men and Asian women, and between Black women and Asians, both males and females. The importance of race was further indicated by an almost

significant (p = .051) negative correlation between the ratings of White women and Black women. To conclude, if correlations between ratings of the different types of faces had been generally positive, an interpretation in terms of response style would be plausible. But since some of the correlations were negative, an interpretation in terms of a race factor is more plausible.

The modern and classical racial prejudice scales were not significantly related to ratings of Asian and White faces. The female Asian faces displayed the same pattern as the Black faces, but weaker.

General Discussion

The pressure to inhibit overtly xenophobic expressions is strong in present-day Western societies (Franco & Maass, 1999; Pettigrew & Meertens, 1995). The violence test gave a direct estimate of this. The instrument could quite legitimately be presented as either a measure of anxiety or a measure of social stereotypes. The effect on both expression and inhibition of the stereotype of the immigrant as aggressive was very strong. The Cohen's d value of 1.19 can be compared to the mean d value of .41 obtained in ten studies on prejudice using the bogus pipeline technique (reported in Roese and Jamieson, 1993). The reports reflected our participants' interpretation of the situation. The anxiety instruction produced an instrument measuring reported affect toward a target where the critical feature is masked (for reviews concerning related instruments, see Fazio & Olson, 2003; Maass, Castelli & Arcuri, 2000; Nosek, 2002). The stereotype instruction produced stereotype compensation efforts.

We expected sex differences to appear in both the implicit anxiety condition and the explicit stereotype condition. In both cases, females generated more prejudiced reports than males. This was predicted from the nature of the stereotype, one of the foreigner and immigrant as dangerous and aggressive. A fear-related appraisal is involved, and women could be expected to score higher than men based on a perceived smaller capacity for handling aggressive situations. Since an ethnic stereotype is involved, the phenomenon could be called *fear-based* racism. Women report lower prejudice than men do on questionnaires in many areas (Ekehammar et al., 2003). Apparently fear-based racism is an exception to this rule.

In Study 1, an attempt was made to disguise the faces test as a test of general attitudes based on first impressions. Given the situation, that faces clearly different in ethnicity are to be rated on valence-related adjectives, the instrument seems hard to disguise. But no participant described the instrument as a measure of racism, stereotypes or the like in the

halo/anxiety condition. Further, the manipulation unexpectedly resulted in participants under the stereotype instructions rating Whites even more positively (as compared to Blacks/Asians) than did participants under the halo instruction. The main finding related to correction was replicated in Study 2. The participants showed a socially desirable way of responding to photographs of Black men, favoring them over all other groups. But they gave low ratings to Black women. The reports in all conditions may be based on attempts to avoid the negative emotions produced by the Black male faces and the threat of being bigoted implied by this. However, the fact that in the present study each ethnic category was represented with a rather limited number of faces calls for caution in interpretation of these results. A replication with a wider sample of exemplars from each category would increase their validity.

In Study 2, the relationship between correction in the faces test and explicit measures of racism and the motivation to control prejudice was investigated. No significant correlations were found between correction and the explicit measures. Weak correlations between explicit and implicit measures are very common in this area (Olson & Fazio, 2003) and may suggest that the correction process was not strategic and deliberate. On the other hand, the significant correlation between explicit racism and ratings of the Black faces indicates that subjects were influenced also by their explicit attitude towards this group. Also, it increases the general validity of the faces test as a measure related to stereotyping and prejudice.

The limitations of working memory have been known for a long time. Our cognitive economy is one of scarcity, so we have to process social stimuli in low-budget ways (e.g., Kunda, 1999). Categorization is a wonderful tool in this regard, allowing great parsimony of representation. This was quite clear to Gordon Allport (1954), who also recognized that social categorization is made in terms of ingroup and outgroups and often involves extremely negative evaluations. When the topic was brought up again around 1990, social psychologists pointed to our capacity for self-correction. If we think that our judgment is biased, we can use correction processes and thus overcome stubborn negative stereotypes.

There are a number of limitations to this optimistic view. First, early phases of information processing can be accepted as reality by later phases. This builds prejudice into our very view of reality (Gawronski, Geschke & Banse, 2003; Hagenberg & Bodenhausen, 2003). Secondly, when our cognitive resources are low or engaged, correction will not be possible, a point stressed by Bargh (1999). Thirdly, monitoring for bias is more or less automatic but correction processes are effortful (Wegner, 1994). This brings us back to the economy of scarcity. Our participants seem to apply a simple correction device, overcompensation, to Black men. But they do not apply correction to their evaluation of

Black women. This is so for almost all our participants and shows limited variation (despite the difference in level, there is still a significant positive correlation between the Black man and Black woman scores). It is neither related to conscious motivation to avoid prejudice, nor to naive theories of attitudes toward different ethnic groups. Instead, it seems to be an attempt to uphold a spotless view of the self, as described in Dovidio and Gaertner's (1998) account of aversive racism.

References

Akrami, N., Ekehammar, B., & Araya, T. (2000). Classical and modern racial prejudice: a study of attitudes toward immigrants in Sweden. *European Journal of Social Psychology*, *30*, 521-532.

Allport, G. W. (1954). *The nature of prejudice*. Garden City, NY: Addison-Wesley.

Araya, T., Akrami, N., Ekehammar, B., & Hedlund, L-E. (2002). Reducing prejudice through priming of control-related words. *Experimental Psychology*, *49*, 222-227.

Banerjee, R. (2002). Children's understanding of self-presentational behavior: Links with mental-state reasoning and the attribution of embarrassment. *Merrill-Palmer Quartlery, 48*, 378-404.

Bargh, J. A. (1999). The congitive monster: The case against controllability of automatic stereotype effects. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 361-382). New York: Guilford.

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

Baumeister, R. F., Dale, K., & Sommer, K. L. (1998). Freudian defense mechanisms and empirical findings in modern social psychology: Reaction formation, projection, displacement, undoing, isolation, sublimation, and denial. *Journal of Personality, 66*, 1081-1124.

Bodenhausen, G. V. & Moreno, K. N. (2000). How do I feel about them? The role of affective reactions in intergroup percpetion. In H. Bless & J. P. Forgas (Eds.), *The message within. The role of subjective experience in social cognition and behavior* (pp. 283-303). Philadelphia, PA: Psychology Press.

Chaiken, S., & Trope, Y. (Eds.), (1999). *Dual-process theories in social psychology*. New York: Guilford Press.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*, 5-18.

Dovidio, J. F., & Gaertner, S. L. (1998). On the nature of contemporary prejudice: The causes, consequences, and challenges of aversive racism. In J. L. Eberhardt & S. T. Fiske (Eds.), *Confronting racism: The problem and the response* (pp. 3-32). London: Sage.

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.

Ekehammar, B., Akrami, N., & Araya, T. (2003). Gender differences in implicit prejudice. *Personality and Individual Differences, 34*, 1509-1523.

Fazio, R. H., & Dunton, B. C. (1997). Categorization by race: The impact of automatic and controlled components of racial prejudice. *Journal of Experimental Social Psychology*, *33*, 451-470.

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

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

Forgas, J. P. (2000). Affect and information processing strategies: An interactive relationship. In J. P. Forgas (Ed.), *Feeling and thinking: The role of affect in social cognition* (pp. 253-282). Cambridge: Cambridge Universitiy Press.

Franco, F. M., & Maass, A. (1999). Intentional control over prejudice: When the choice of the measure matters. *European Journal of Social Psychology*, *29*, 469-477.

Gawronski, B., Geschke, D., & Banse, R. (2003). Implicit bias in impression formation: associations influence the construal of individuating information. *European Journal of Social Psychology*, *33*, 573-589.

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

Guimond, S., Dambrun, M., Michinov, N., & Duarte, S. (2003). Does social dominance generate prejudice? Integrating individual and contextual determinants of intergroup cognitions. *Journal of Personality and Social Psychology*, *84*, 697-721.

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

Kinder, D. R. (1986). The continuing American dilemma: White resistance to racial change 40 years after Myrdal. *Journal of Social Issues*, *42*, 151-171.

Kinder, D. R., & Sears, D. O. (1981). Prejudice and politics: Symbolic racism versus racial threats to the good life. *Journal of Personality and Social Psychology*, *40*, 414-431.

Kunda, Z. (1999). *Social cognition: Making sense of people*. Cambridge, MA: The MIT Press.

Maass, A., Castelli, L., & Arcuri, L. (2000). Measuring prejudice: Implicit versus explicit techniques. In D. Capozza & R. Brown (Eds.), *Social identity processes* (pp. 96-116). London: Sage.

McConahay, J. B. (1986). Modern racism, ambivalence, and the modern racism scale. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 91-126). New York: Academic Press.

Millham, J., & Kellogg, R. W. (1980). Need for social appraisal: Impression management or self-deception? *Journal of Research in Personality, 14*, 445-457.

Nosek, B. A. (2002). *Moderators of the relationship between implicit and explicit attitudes*. Unpublished Ph.D. dissertation, Yale University.

Paulhus, D. L., & Reid, D. B. (1991). Enhancement and denial in socially desirable responding. *Journal of Personality and Social Psychology*, *60*, 307-317.

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

Roese, N. J., & Jamieson, D. W. (1993). Twenty years of bogus pipeline research: A critical review and meta-analysis. *Psychological Bulletin, 114*, 363-375.

Schuman, H., Steeh, C., & Bobo, L. (1985). *Racial attitudes in America: Trends and interpretations*. Cambridge, MA: Harvard University Press.

Sears, D. O. (1988). Symbolic racism. In P. A. Katz & D. A. Taylor (Eds.), *Eliminating racism: Profiles in controversy* (pp. 53-84). New York: Plenum.

Sidanius, J., & Pratto, F. (1999). *Social dominance*. Cambridge, MA: Cambridge University Press.

Sniderman, P. M., & Tetlock, P. E. (1986a). Symbolic racism: problems of motive attribution in political analysis. *Journal of Social Issues*, *42*, 129-150.

Sniderman, P. M., & Tetlock, P. E. (1986b). Reflections on American racism. *Journal of Social Issues*, *42*, 173-187.

von Hippel, W., Silver, L. A., & Lynch, M. E. (2000). Stereotyping against your will: The role of inhibitory ability in stereotyping and prejudice among the elderly. *Personality and Social Psychology Bulletin, 26*, 523-532.

Wegener, D. T., & Petty, R. E. (1995). Flexible correction processes in social judgment: The role of naive theories in corrections for perceived bias. *Journal of Personality and Social Psychology*, *68*, 36-51.

Wegner, D. M. (1994). Ironic processes of mental control. *Psychological Review*, *101*, 34-52.

Appendix A

The adjective pairs used in the faces and the violence tests (Swedish words in parenthesis)

Faces test

Calm (Lugn)	Uneasy (Ängslig)
Hostile (Fientlig)	Friendly (Vänlig)
Social (Sällskaplig)	Reclusive (Enstörig)
Passive (Passiv)	Outgoing (Utåtriktad)
Reliable (Pålitlig)	Unreliable (Opålitlig)
Callous (Hårdhjärtad)	Tender-hearted (Ömsint)
Competent (Kompetent)	Ignorant (Okunnig)
Untalented (Obegåvad)	Intelligent (Intelligent)
Sluggish (Trög)	Lively (Livfull)
Cheerful (Glad)	Dejected (Nedstämd)

Violence test

Tense (Spänd)	Quiet (Lugn)
Sad (Sorglig)	Cheerful (Glad)
Closed (Sluten)	Open (Öppen)
Resigned (Resignerad)	Confident (Hoppfull)
Relaxed (Avspänd)	Threatening (Hotfull)
Pleasant (Trevlig)	Unpleasant (Otrevlig)
Agreeable (Behaglig)	Uncomfortable (Olustig)
Dull (Trög)	Vehement (Häftig)
Friendly (Vänlig)	Angry (Ilsken)
Violent (Våldsam)	Calm (Stilla)