



SONDERFORSCHUNGSBEREICH 504

Rationalitätskonzepte,
Entscheidungsverhalten und
ökonomische Modellierung

No. 04-39

**Determinants and Consequences of Survey
Respondents' Social Desirability Beliefs about
Racial Attitudes**

Volker Stocke*

October 2004

Stimulating discussions with Hartmut Esser, Herbert Bless and Stephan Ganter are gratefully acknowledged. Tobias Stark and Diana Schirowski were a great help in preparing this article. This research was supported by a grant of the German Science Foundation (DFG) to the Sonderforschungsbereich 504 at the University Mannheim.

*Sonderforschungsbereich 504, email: vstocke@rumms.uni-mannheim.de



Universität Mannheim
L 13,15
68131 Mannheim

Abstract

In this article we analyze beliefs about the social desirability of ten racial attitude items from the German General Social Survey (ALLBUS). These beliefs indicate that the items, as well as respondents with regard to different sex, age and education, are differently prone to social desirability bias. Demographic response differences may thus only reflect differences in social desirability bias. We matched the desirability differences between the items and demographic groups with the racial attitude responses from the independent, nationwide sample of the ALLBUS survey in 1996. The desirability beliefs obtained from our urban, West German sample predicted the attitude answers, and this predictability was stronger for ALLBUS respondents with the same characteristics. Our results suggest that the ALLBUS data is subject to social desirability bias, that particular items are more strongly affected, and that differences in the attitude reports according to the respondents' age and education should be interpreted with caution.

Keywords: desirability beliefs, racial attitudes, social desirability bias, trait desirability.

Introduction

Social desirability bias (SD-bias) can seriously affect the validity of survey data and the appropriateness of hereon based research. For instance, respondents have been found to underreport their substance use, as well as the prevalence of symptoms of mental illness, and to have over-reported the frequency of attending religious services (Epstein, Barker, & Kroutil, 2001; Presser & Stinson, 1998). Because of their systematic nature, social desirability effects always bias univariate response distributions, and thus survey results do not represent the real prevalence of the analyzed characteristics in the sample. One could however argue that biased response distributions are not a serious problem when researchers are primarily interested in the nature of the associations between the attribute under consideration and other characteristics of the respondents, as for example their sex, age or education. These associations will however only remain unaffected by social desirability when respondents with different characteristics are subject to bias of a similar strength and direction. When instead, respondents with characteristics of interest as explanatory factors are subject to systematically different SD-bias on the outcome dimension, this may create artificial or suppress real associations. When social desirability effects bias the response distributions and in particular the structure of observable associations, survey researcher may wish to reduce these effects by introducing appropriate statistical controls when analyzing the data. This is however only a promising strategy when the strength as well as the direction of SD-bias, and in particular possible group differences in this respect, are well known.

The social desirability of those characteristics which respondents ascribe to themselves, when they select a certain answer in self-description questionnaires, is an important, but surprisingly little analyzed determinant for the strength and direction of SD-bias (cf. Edwards, 1957). The trait desirability respondents perceive in an interview situation is an important factor, which shapes their expectations about evaluative reactions from others, when they answer a survey question. The desirability beliefs are the respondents' perceptions of the trait desirability and at the cognitive core of all theoretical approaches in which SD-bias is assumed to result from respondents' impression management strategies. In these approaches, the subjects aim is to reach positive evaluations from others. The Expected Value Theory of response behavior is a theoretical framework in which the approval from others is assumed to be the motivation underlying social desirability effects (Esser, 1991; Stocké, 2004a; Tourangeau, Rips, & Rasinski, 2000, p. 281). In order to form expectations about which answer is best suited to gain approval and to avoid disapproval, respondents utilize their subjec-

tive desirability beliefs: the more they assume the desirability of available response options to differ, the more incentives for SD-bias are perceived.

The few available studies have indeed shown for different survey topics that the perceived desirability of a particular trait predicts the probability that respondents endorse the respective response option in survey interviews (cf. for instance: Nicotera, 1996). There are however doubts about the external validity of these results. In nearly all of these studies, the desirability beliefs and the response behavior about the survey topic under consideration were recorded in the same interview. This leaves the possibility open that thinking about the trait desirability activated the respondents' social desirability concerns, which may not have influenced the subjects' response behavior otherwise. In this case, the SD-bias and the hereon based association between the desirability beliefs and the response behavior must be regarded as an artifact of the research procedure. Whether the perceived trait desirability predicts the response behavior and thus the existence of SD-bias when both kinds of data are recorded with different samples of respondents, has not yet been tested.

Socially desirable response behavior is the possible consequence when the desirability of the response options are perceived to differ and the respondents' 'true' characteristics do not conform with the most positively evaluated option. This may often be the case in the field of racial attitudes and racial prejudice. Empirical evidence suggests that answers about this topic are subject to SD-bias. Studies from the U.S. have tested whether white respondents' reports about their attitudes toward Afro-Americans differs between interviewer- and self-administered surveys (Krysan, 1998). Across all items the racial attitude answers were found on average to be significantly more negative when the privacy of responses was guaranteed under self-administration. However, the strength of this privacy effect differed considerably between the specific items. This suggests that the respondents' desirability beliefs may differ between the attitude questions. Whether this is the case, has not been tested empirically. Furthermore, some evidence indicates that the desirability beliefs not only varies between different racial attitude questions, but according to the socio-demographic characteristics as well (Reinecke, 1991, p. 152). Whether these differences in the desirability beliefs affects the racial attitude answers was not tested in this study.

The respondents' perceptions of social norms are the basis for their beliefs about how desirable traits are regarded in society. Although norms are in principle socially shared, they may differ considerably between groups of respondents. This is the case, since social norms are often more binding or even prescribe different traits for subjects with particular group affiliations. Such a situation has been found for subjects with different genders, in the case of

norms regulating sexual promiscuity. Here, men were found to over-report and women to under-report the number of sex partners (Smith, 1992). Furthermore, the perception of social norms underlying the respondents' desirability beliefs are likely to be shaped by their embeddedness in and identification with particular reference groups, and thus may differ according to the composition of this social context (Singer, 1981). Only little empirical evidence about group differences in the desirability beliefs in the field of racial attitudes is available. Although differences in the SD-bias between socio-demographic groups are in particular negative for the validity of survey data, no research has been done about whether such differences have to be expected in survey data about racial attitudes.

The present study has three aims. *First*, to analyze in detail the strength and direction of desirability beliefs about racial attitude answers on 10 items from the German General Social Survey (ALLBUS). We show whether these beliefs, and thus the incentives for SD-bias, differs between the items and groups of respondents. The *second* aim is to provide evidence for whether these differences in the perceived desirability predicts attitude responses collected in a survey with an independent sample. For this purpose we combined the desirability beliefs from our own survey with the response behavior in the nationwide ALLBUS survey from 1996. Since the desirability beliefs were obtained from West German respondents, living in a urban context, we expect these beliefs to be more predictive for the responses of the same subset of respondents in the ALLBUS survey.

Theoretical Background

In the expected value theory of response behavior, the answers of survey participants and the decision to respond in a socially desirable manner are explained as the result of a goal directed and instrumentally rational selection between response options (Esser, 1991; Stocké, 2004a; Tourangeau et al., 2000, p. 281). From this perspective, the respondents' aim when they answer a self-descriptive questionnaire is to achieve positive and avoid negative reactions from others, and thus to maximize their feeling of approval. However, the precondition that subjects expect their answers to be instrumental in this respect is that others are both present in the response situation *and* able to observe the answers.¹ Furthermore, it is of uppermost importance for the respondents to anticipate whether and of which sort a certain answer will provoke evaluative reactions from others present in the interview situation (Esser, 1991; Stocké, 2001). The basis for these expectations, and thus the cognitive determinant for predicting the strength and direction of SD-bias, are the respondents' desirability beliefs. These beliefs are about whether and which evaluative reaction have to be expected from other per-

sons present in a specific interview situation and when a certain answer is selected on a particular question. Respondents can be assumed to utilize different kinds of information in order to arrive at their desirability beliefs, even when, as is typically the case for the unknown interviewer, individualized prior knowledge about the evaluation criteria is not available. Indirect evidence about the most likely reactions may be provided by the interviewers' reactions to previous answers or their visible characteristics, which the respondents perceive to be associated with particular evaluation criteria. When this information is not available, respondents rely on their knowledge about the general social desirability of the respective trait in society (Stocké, 2004a). These beliefs are the basis for the respondents' best guess about the most favorably evaluated answer in the particular interview situation.

From the aforementioned theoretical perspective, the SD-Bias is a function of the strength of the respondents' approval motive and to what degree response privacy is absent. Furthermore, and in the center of our present study, the strength *and* direction of SD-bias results from the respondents' desirability beliefs. Thus, subjects are attracted to the direction of a particular response category, the more the desirability of this category is perceived to be superior to other available options. In summary, the total motivation for socially desirable response behavior is determined by the three factors mentioned above, in which each is a necessary, but not sufficient determinant for SD-bias: a strong approval motive makes evaluation reactions for the respondents relevant, subjects regard such reactions as possible because of insufficient privacy, and desirability beliefs are such that choosing one or another option makes a difference. If only one of these conditions is not given, the effect of all other factors on the probability of SD-bias is assumed to vanish and subjects are expected to answer according to their subjectively true scores. The total incentives for socially desirable response behavior can be represented as follows:

$$SEU(SD) = U_{SD} \bullet w_P \bullet \Delta W_{TD}$$

The *first* parameter, U_{SD} , represents the strength of the approval motive. This parameter has a value of one for subjects with a strong approval motive and zero when this motive is absent. The *second* parameter, w_P , stands for perceptions about how likely the answers in the particular response situation can affect the satisfaction of the approval motive. This parameter is zero under the condition of complete privacy and one when others are able to perceive the answers. The *third* parameter, ΔW_{TD} , represents the subjects' desirability beliefs, or more precisely, the perceived desirability differential between the response options. This parameter varies between -1 and +1, depending on which option is regarded as being more desirable.

Empirical Evidence from Previous Research

Racial attitude reports and social desirability bias

Different methods have been used to test whether, and in which direction, respondents' racial attitude reports are subject to SD-Bias. One line of research has analyzed the association between the respondents' attitude reports and their need for social approval, as measured with different social desirability (SD)-scales. Subjects with high SD-scale values are assumed to distort their responses more in the direction of social desirability, and are expected to report those answers they regard as socially desirable (Crowne & Marlowe, 1964). Thus, correlations between SD-scale scores and attitude responses are taken as evidence for the existence of SD-bias. Results from this research are mixed. Mielke (1995) asked the respondents about their racial attitudes using Pettigrew and Meertens' Subtle and Blatant Prejudice scale. Findings have shown that responses on both instruments were more in the direction of liberal racial attitudes when respondents scores on a SD-scale indicated a stronger need for social approval. Duck and Hunsberger (1999), found the answers on the 'Manitoba Prejudice' scale, which measures prejudice against immigrants, not to be correlated with the subjects' scores on the Marlow-Crowne SD-scale.

Other research has utilized the bogus pipeline paradigm to test for the presence of SD-bias. Here, the respondents were led to believe that the researcher had a physiological measure for honest responding. This is reached with seemingly complex equipment, similar to a lie detector. In the pioneer study, white American respondents rated whether positive and negative adjectives are appropriate characterization for white and black persons (Sigall & Page, 1971). A random half of the sample was tested under bogus pipeline conditions and the other half under standard interview conditions. Accordingly, subjects reported more negative evaluations about blacks when they believed that the honesty of their answers could be checked. In another study the respondents answered 24 items supposed to measure prejudice against Turks either under a standard interview or a bogus pipeline condition (Wagner & Zick, 1995). The respondents reported more negative racial attitudes under bogus pipeline conditions. These effects were replicated in other studies as well (Mummendey, Bolten, & Isermann-Gerke, 1982, Schlenker, Bonoma, Hutchinson, & Burns, 1976).

Another, more indirect method for finding the effects of social desirability bias is to utilize the non-reactive priming and response latency procedure to measure racial attitudes (Fazio, Jackson, Dunton, & Williams, 1995). This procedure is based on the subjects' uncontrolled evaluative reactions when a typical member of the respective ethnic outgroup is presented. Whether a positive, a negative or no evaluation is spontaneously activated is measured

with the response time the subjects need to identify positive or negative stimuli, after the activation sequence. Racial attitudes measured in such a non-reactive way were found to be on average more negative compared with the subjects' self-reports. Since the priming and response latency measure is obtained without the subjects' awareness and is therefore regarded to be unaffected by conscious impression management strategies, the differences in the self-reported measure were assumed to be the consequence of SD-bias.

Two studies from the U.S. tested whether white respondents' reports about their attitudes toward Afro-Americans differs between interviewer- and self-administration (Krysan, Schuman, Scott & Beatty 1994; Krysan, 1998). It was assumed that responses under self-administration, and thus ensured privacy, are more in line with the respondents' 'true scores', compared with the interviewer-administered condition. In the second case, the interviewer is able to perceive the responses, and is thus the target of self-presentation strategies. On average, across all items used in these studies, racial attitude answers were significantly more negative under the condition of guaranteed privacy. On the level of the single items, the strength and partly the direction of the privacy effects varied considerably: for nine of the nineteen attitude questions, private attitudes were significantly more negative, for eight items there was a non-significant tendency in this direction and for two items a tendency in the opposite direction. This may be regarded as evidence that the respondents had desirability beliefs of different strength and direction in terms of the specific item content.

Operationalization of desirability beliefs

How respondents perceive the general evaluation of a certain trait in society can be measured in the form of 'personal' or as 'extrinsic expected' desirability. In the first case, subjects report their personal evaluation of a person who possesses the respective characteristic (cf. Meleddu & Guicciardi, 1998). The respondents' aggregated evaluations on this measure are an indicator for the normative climate in the respective sample, and the population this sample represents. This is a valid operationalization for those informal sanctions a survey respondent expects when revealing the respective trait, if the objective distribution of judgments is in accordance with the respondents' subjective perceptions. Empirical research has shown that this is at least not always the case (Crott & Roßrucker, 1974). The second operationalization of the trait desirability measures directly those evaluations which are believed to be dominant in society (Edwards, 1957; Nicotera, 1996). Respondents report those evaluation reactions they expect from an average person when a particular characteristic is revealed.

In the case of gradually varying traits, as for example being more or less prejudiced, the researcher barely always measured the desirability of the presence of the respective trait. The respondents judge the desirability of the trait on a bipolar response scale with positive and negative endpoints, the midpoint of the scale representing a neutral evaluation. Incentives for SD-bias are assumed to increase, the more the judged desirability deviates from this midpoint. This measure however assumes that an absence of the evaluated characteristics is neutrally evaluated. When this does not apply, the real incentives from social desirability are either over- or underestimated (Stocké & Hunkler, 2004). Thus, an appropriate operationalization of the incentives for SD-Bias necessarily uses the difference between the evaluation of a strong presence and a complete absence of the respective trait (cf. Stocké, 2004b).

Another issue when predicting the incentives for SD-bias is the appropriate level of data aggregation. Desirability beliefs have been mostly analyzed and reported on the aggregate level of the sample. This is only an appropriate measure for the potential SD-bias, when the direction of perceived incentives is homogeneous in the sample. When this is not the case, differences between individuals and subgroups of respondents cancel each other out on the aggregate level, and the real strength of SD-bias may be largely underestimated. An analysis of desirability beliefs must thus take a possible heterogeneity in this respect into account.

Evidence for the explanatory power of desirability beliefs

All studies we are aware of confirmed the hypothesis that respondents' beliefs about the desirability of a certain trait predict the probability that the respective response option is endorsed in a self-description questionnaire. Gove and Geerken (1977) for instance have shown in an early study from this field of research, that differences between subjects as to how desirable they regard high self-esteem or a positive affect, predicted the probability that the respective response options were endorsed. This was replicated in a study from Phillips and Clancy (1972) with seven sensitive survey topics and by Nicotera (1996) with the 20 items from the Infante and Rancer's Argumentativeness Scale. Another study analyzed the association between beliefs about the desirability of police arrests and the self-reports about this topic (Wyner, 1980). Since the true number of arrests was obtained from police records, it was not only possible to analyze the correlation between responses and desirability beliefs, but to test the predictive power of these beliefs for the strength and direction of SD-Bias. Results have shown that desirability beliefs predict whether and how strongly subjects under- and over-report the number of their police arrests in the past.

In a study by Phillips and Clancy (1972), subjects rated the social desirability of being unprejudiced against ethnic outgroups. Subjects reported in the same interview on how bothered they were by meeting persons from other ethnic groups. Respondents who regarded an unprejudiced attitude as being highly desirable reported more positive attitudes than those with medium and low desirability ratings. In a German study, respondents rated the social desirability of 37 racial attitude items (Reinecke, 1991, p. 152). On average, subjects judged positive responses to be more desirable. The desirability judgments, however, differed significantly according to the respondents' age, as well as their socioeconomic status and varied between the different attitude questions. Whether the desirability beliefs explained the response behavior on the respective attitude items was however not tested. In a study from Stocké (2004b) it was analyzed whether the respondents' need for social approval, their desirability beliefs and the privacy of the response situation in combination affects the respondents' racial attitude answers. The results show that desirability beliefs determine how strong, and in which direction the responses were affected, when others were able to observe the response behavior. This interaction effect between the desirability beliefs and the response privacy was more pronounced for subjects with a stronger approval motive. However, since the respondents' desirability beliefs and their attitude answers were collected in the same interview, the external validity of these results is an open question.

In a study from Huang and colleagues (1998), the possible problem of an artificial activation of social desirability concerns when subjects were asked about their characteristics and the judged desirability of these characteristics in the same interview, has been attenuated. Here, the social desirability ratings of 288 items in testing subjects' mental health were recorded in a first interview. In a second session, the same subjects responded to these items. Despite the time distance between both answers, subjects reported fewer symptoms of mental illness when they judged this trait to be less desirable. In a study by Chen and colleagues (1997), respondents initially rated the 45 items of the PANAS positive/negative affectivity scale with respect to social desirability. In the second part of the study, members of an independent sample were asked whether the personality traits described by the PANAS items applied to themselves. Subjects' approval motive was measured using the Marlowe-Crowne SD-scale. As a result, the on the level of the items aggregated desirability beliefs from the first study substantially predicted the responses in the second study. This correlation was significantly stronger for subjects with a strong approval motive.

Empirical Study

Sample and data collection

The respondents in our *trait desirability study* were a multi-stage, local random probability sample of residents from a metropolitan area in the western part of Germany. This community has about 480.000 inhabitants. Households were listed with a random walk procedure and respondents selected using the 'last-birthday' method. Because the topic of the study was the beliefs about the desirability of racial attitudes, we only included subjects with a German citizenship in the sample. The 378 interviews took place in the respondents' homes and were realized interviewer-administrated with paper and pencil questionnaires. The response rate was 70.7 percent. The response behavior about racial attitudes was taken from the German General Social Survey (ALLBUS) conducted 1996. This survey used a representative random probability sample of the East and West German population. The 3.518 interviews were realized interviewer-administered with paper and pencil questionnaires at the respondents' homes. The response rate was 54.2 percent. In the sample, 212 respondents, this are 6.0 percent, did not have a German citizenship and were thus excluded from the analysis. Furthermore, 77 respondents did not answer all questions about their racial attitudes. Thus 3.229 cases were left for our analysis, in which 2.137 respondents, 66.2 percent, were from the western part of Germany and 1.092 respondents were interviewed in the eastern part of the country. In this sample, 33.8 percent of the respondents lived in cities with more than 100.000 inhabitants, whereas the other 68.5 percent came from smaller cities or villages.

Measures

- *Racial attitude responses*: The 10 racial attitude items, which were included in our study, were utilized in the 1996 study of the German General Social Survey (cf. the item wording in table 1).² The respondents' answers were recorded using seven-point Likert scales with endpoints labeled with 'completely agree' and 'completely disagree'. For our analysis, the scoring of all items was recoded in such a way that low attitude scores represent negative and high scores positive racial attitude answers.

- *Desirability beliefs and the relative desirability of positive and negative racial attitudes*: The desirability beliefs were operationalized as 'extrinsic' evaluations, that the respondents anticipated as a typical reaction from others when certain racial attitudes were revealed to an unknown counterpart (Edwards, 1957). Respondents were asked to imagine a situation such as a train ride, in which a conversation between two strangers takes place. This scenario was used in order to mimic a typical interview situation. Subjects reported their anticipation

of how embarrassing it would be for one of these persons to disclose certain opinions about foreigners (cf. for the exact question wording: Stocké 2004b). The altogether 20 opinions judged in this way were identical to those expressed, when either completely disagreeing or agreeing with the 10 racial attitude items used in our study. The desirability beliefs were recorded with a bipolar response scale from -4 (this statement would be very embarrassing) to +4 (this statement would be very pleasant). After the responses were recoded into a range of values between 0 and 8, for each respondent and attitude item the judged desirability of a negative racial attitude statement was subtracted from that of a positive attitude. The relative desirability scores obtained ranges from -8 (negative racial attitudes more desirable) to +8 (positive attitudes more desirable). A value of zero indicates that positive and negative attitudes were regarded as being equally evaluated in society.

- *Differentiation of desirability beliefs in the population:* Desirability beliefs are predictive for the incentives of SD-bias, when the informants and the subjects who answer the racial attitude questions share the same perception of social norms in society. This has however been shown not to be the case. For instance, men were found to evaluate a crafty person as more socially desirable than women (Mecklenbräuker, Hager, & Möller, 1994). Cultural differences between countries have been found as well. Thus, persons being 'aggressive', 'self-seeking' and 'earnest' were found to be less negatively evaluated by Americans than by British respondents (Hampson, Goldberg, & John, 1987). One important determinant for cultural diversity and normative differences within countries are regional variations in urbanity. Accordingly, social values have been found to differ on a variety of dimensions between rural and urban areas in Germany (Böltken, 1992). The most pervasive cultural division in Germany is however still the difference between the region of the former communist East Germany and the western part of the country. Empirical research has found considerably distinctive social norms and values in these two parts of the country. For instance, in West compared with East Germany, sex roles are more traditional, norms against unconventional political protest are stronger and tolerance for minority groups is judged to be more desirable (Fuchs, 2000; Herbert, 1991). We therefore expect the desirability beliefs of respondents from areas of different urbanity, as well as those from East and West Germany to differ. Thus, beliefs recorded in our sample from an urban, West German region should predict less the SD-bias in a survey with respondents from a rural and East German area.

In the ALLBUS study from 1996, respondents were asked about how they expect the majority of Germans would evaluate certain discriminating behavior against foreigners. The scenarios are: a foreigner is not served in bar, parents don't allow their daughter to have a

foreigner as a friend, and an employer first fires non-German employees. Answers on these questions, recorded on a four-point response scale and higher values expressing stronger beliefs, can be assumed to reflect the perceived normative climate against racial discrimination in the country. Since we are not aware of any study which has analyzed regional differences in the desirability beliefs about racial attitudes, we computed the mean perceived normative climate against the three forms of racial discrimination. We found for West Germans a score of 2.5 and for East Germans a score of 2.3. Thus subjects in the western part of the country perceived significantly stronger norms against racial discrimination ($t=8.1$, $df=3449$, $p \leq 0.05$). The same comparison between respondents from communities with more and less than 100.000 inhabitants found a score of 2.5 for the more urban and of 2.4 for the more rural regions. Thus, respondents from more urban areas perceive slightly, but statistically significant stronger norms against discrimination ($t=2.9$, $df=3449$, $p \leq 0.05$).

Results

Sample characteristics

In table 1, the respondents' from the trait desirability and the ALLBUS studies are compared with respect to their demographic characteristics. We found significant differences in the age and education between the two samples, which are likely to be the result of the different sample frames and response rates. Thus, the subjects in the ALLBUS study were on average 2.9 years older than those in our trait desirability study ($t=3.1$, $df=3605$, $p \leq 0.05$). Furthermore, subjects with compulsory education were 7.3 and those with university degrees 2.3 percentage points more represented than in the trait desirability study. This leads respondents with a high school certificate to be 6.1 and those with a secondary school certificate 3.5 percentage points less included ($\text{Chi}^2=17.5$, $df=3$, $p \leq 0.05$). The samples did not differ significantly according to the sex of the respondents ($\text{Chi}^2=0.1$, $df=1$, $p > 0.05$).

- table 1 about here -

Descriptive analysis of desirability beliefs

In our sample, respondents in general expected that positive racial attitude answers are favorably evaluated in society, but how favorable varied between the attitude items (cf. table 2, column 1). The aggregated judgments on the response scale from +4 (very desirable) to -4 (very undesirable) ranged between +0.08 for item 2 and +0.95 for item 9. On average across the ten questions this score is +0.57. Answers which expressed negative attitudes toward for-

eigners received ratings between +0.52 in the case of item 1 and -0.70 in the case of item 8 (cf. table 2, column 2). The mean desirability score of -0.34 across all attitude questions indicated that respondents judged negative racial attitudes to be in general unfavorably evaluated in society. In column 3, the aggregated relative desirability scores, this is the difference in the anticipated evaluation of positive and negative racial attitude answers, summarizes the incentives for SD-bias which is predicted for each of the attitude items. For nine out of the ten attitude items, a positive relative desirability score indicated that respondents on average perceived incentives to give positive racial attitude answers.

- table 2 about here -

When the desirability beliefs in the sample are heterogeneous with respect to the direction of incentives from social desirability, indices aggregated on the item level do not reflect the true incentives for SD-bias. In order to test whether this applies in the case of our racial attitude items, we analyzed the relative desirability beliefs in table 3, differentiated according to the sign of the individual judgments. The results clearly indicated that respondents do not agree about whether positive or negative racial attitude answers are more desirable in society: on average across all items a majority of 53.7 percent assume positive attitudes to be more desirable, but 24.6 percent assumed this for negative attitudes and 21.7 percent did not perceive any desirability differences. With 64.8 percent, the highest proportion of respondents perceived incentives for a positive attitude answer in the case of agreeing with item 10, thus saying that foreigners are an enrichment for German culture. In the case of item 1, only 34.4 percent and thus nearly half so many subjects perceived a positive attitude answer to be the more desirable response option. The percentage of respondents, who assumed negative attitude answers to be more desirable, differed considerably between the different attitude questions as well: here, 40.5 percent regard it to be more desirable to deny the right of foreigners to receive the same welfare benefits than the Germans (item 2). In the case of items 9 and 10, only 14.5 percent believed that a negative attitude answer, that is to deny the value of foreigners for the culture in Germany and the opinion that because of the presence of foreigners one feels as a stranger in Germany, is the socially most desirable answer. On average across the entire attitude scale, respondents with positive desirability beliefs perceived incentives of 2.8, and those with negative beliefs 2.4 points on the relative desirability scale. Since now the desirability beliefs with different signs do not cancel each other out on the aggregate level, we found substantially higher incentives for SD-bias in our individual level analysis.

- table 3 about here -

We calculated for each item the average absolute value of the individual respondents' relative desirability beliefs, including those subjects who do not perceive any desirability differences (cf. the last column in table 3). This indicator for the total potential of SD-bias indicated that for item 5, this is the question about whether respondents believe that foreigners are a burden for the social security system, has a score of 2.5 and is thus potentially strongest affected by social desirability bias. The question of whether foreigners are regarded as a cause of problems in the housing market (item 4) is perceived to be the least sensible question: here, the total incentives for SD-bias is only 1.6. Thus the strength of potential SD-bias varies strongly between the different attitude questions.

- table 4 about here -

In order to test significant differences in the perceived total incentives for socially desirable responding between the ten racial attitude items, we performed t-tests for each pair of attitude questions. The results of these tests, presented in table 4, indicated that the aggregated relative desirability beliefs differ in 24 out of 45 comparisons significantly between the analyzed items. Another important result visible in table 4 is the fact that the ordering of the items with respect to their potential SD-bias has changed, compared with the analysis on the aggregate level. The items were presented in an order of increasing incentive strength, as it was presented in table 2, column 3. Thus, adjacent items were found to have similar potential for SD-bias in the aggregated analysis. Attitude questions without significant differences in desirability beliefs should therefore all be located near the bisecting line of the matrix. According to our more differentiated analysis this not the case: the pairs of items with non-significant differences, these are marked with gray, are spread over the whole matrix. The correlation between the aggregate and individual level indicator for the predicted SD-bias for the attitude item is only 0.35, and did not reach statistical significance ($p > 0.05$).

Response behavior in the General Social Survey

In the nationwide ALLBUS study in 1996, the average racial attitude responses across all ten items was 4.1. This is nearly identical with the neutral midpoint of the seven point response scale, where higher values indicate more positive attitudes (cf. the last column in table 5). For the different attitude questions, the response scores varied between a relatively negative value of 3.1 (item 1) and a rather positive score of 5.1 (item 8). With respect to the two dimensions, expected to cause differences in the desirability beliefs, we found small but consistent differences in response behavior. Accordingly, respondents living in the western part of Germany or in larger communities with more than 100.000 inhabitants gave more positive

attitude answers, compared with those in the eastern part of the country and who lived in smaller communities (cf. column 1 and 2 in table 5). Respondents from western Germany answered on average the racial attitude questions 0.3 and those from bigger cities 0.4 scale points more positive than the respective complementary group.

- table 5 about here -

Differences in the desirability beliefs between socioeconomic groups

In this section we analyze differences in the respondents' desirability beliefs according to their affiliation with socio-demographic groups. In regression model 1, the respondents' relative desirability beliefs about all 10 attitude items are regressed on their age, sex and education (cf. table 6). Furthermore, differences in beliefs between the items are held constant with item dummies. Since in this analysis the observations are not independent, but clustered within the respondents, we calculated a simple two-level regression model, with the subjects' responses as a first and the respondents as a second level of the analysis. This analysis was done with the STATA module 'xtreg'.

- table 6 about here -

The first result from regression model 1 is that respondents' beliefs about the strength and direction of desirability differences between positive and negative racial attitude answers differ according to their age, sex and their educational degrees. The respondents' age did not prove to have a linear relationship with their desirability beliefs. Instead, respondents up to the age of 50 were found to assume that positive racial attitude answers are more desirable in society, compared with older subjects. Furthermore, males have significantly more negative desirability beliefs compared with females. On the third dimension with significant differences in the desirability beliefs, the respondents' education, subjects with compulsory education and even more so those with a secondary school certificate perceive positive racial attitudes less desirably than persons with a university degree and a high school certificate.

In regression model 2, those age and educational groups were combined, which proved to have similar desirability beliefs in regression analysis 1. On the one hand, we then found a significant contrast between the group of the 18 to 50 year olds and older subjects. And on the other, subjects with less than a secondary school certificate differed significantly from those with more education. On the basis of the regression model 2, we calculated the predicted desirability beliefs of the different socio-demographic groups, holding constant the other factors, which were included in the analysis.³ Accordingly, we found for respondents who are 18 to 50 years old a predicted relative desirability belief of +0.68 and for the older subjects much

weaker incentives for positive racial attitude answers of +.22. In the case of gender groups, the desirability beliefs showed to have different signs: men perceived incentives for negative attitude reports of -.10, whereas women showed a desirability score of +.22, and thus assumed positive racial attitudes to be more desirable. In the case of the groups with different education, less educated respondents perceived a negative desirability score of -.45 and those with more education a positive score of +.22. We can thus conclude that the socio-demographic groups perceived incentives for socially desirable response behavior of considerably different strength and partly even in different directions.

Effects of desirability beliefs on the response behavior in the General Social Survey

In this part of our empirical analysis it is tested whether the responses in the ALLBUS study in 1996 on the same racial attitude items as analyzed in the trait desirability study can be predicted with the relative desirability beliefs reported above. In the first step, we reorganized the response data in a way that each attitude answer represented one observation in the dataset. Since each of the 3,229 respondents answered all 10 attitude items, a total of 32,290 observations were included in the following analysis. In the next step, we utilized the regression equation from model 2, which represents the significant differences in the desirability beliefs, to compute predicted desirability scores for each attitude item and each combination of the respondents' gender, age and education. These predicted values were matched with the response data on the equivalent items and from respondents with the respective combination of characteristics. As a result, the desirability beliefs predicted for the ALLBUS sample had a mean of 0.84 and thus were very similar to the sample average of 0.91 observed in the trait desirability study. As a consequence of the matching procedure, the standard deviation of the desirability beliefs was 0.70 and thus substantially lower than the value of 2.64 in the trait desirability study. The maximum of the desirability scores predicted for the ALLBUS sample was +2.4 and the minimum -0.8. This measure predicted for 13.1 percent of the cases incentives for negative and for 86.9 percent incentives for positive racial attitude answers.

Because of the way our data was organized, the observations partly stem from the same respondents, and thus are not independent. As in the previous analysis, we calculated a two-level regression model in order to correct this problem. With regression model 3, we tested first whether the attitude responses differed between the items and the socio-demographic groups (cf. table 7). The results have shown that compared with the first item, the respondents gave significantly more positive answers on all other attitude questions. Furthermore, respondents older than 50 years and males reported significantly more negative

racial attitudes, compared with younger and female subjects. Furthermore, respondents with a university degree answered the attitude questions significantly more positively than all other groups with less education. These differences in response behavior closely matches the desirability beliefs we found in the previous section.

In regression model 4, we included the matched relative desirability beliefs into the analysis (cf. table 7). The results were as expected. When positive attitude answers are predicted to be perceived increasingly more desirable than negative responses, this influenced the respondents to give more favorable judgments about foreigners. Thus, the differences in the desirability beliefs between the attitude questions and demographic groups of respondents, which were reported from a completely different sample, predict the response behavior in the ALLBUS survey. When the incentives for socially desirable response behavior are included into the analysis, 13 out of 15 of the regression parameter, representing differences in the attitude responses between the items and the demographic groups, were found to be considerably reduced. After controlling for beliefs about the social desirability of racial attitudes, in 9 of these cases the parameters are even non-significant. The only exceptions were the contrast in the response behavior between respondents with a high school certificate and compulsory education, which remained stable, and the gender differences, which were found to be slightly strengthened. The effect of the respondents' age and education on their attitude responses was strongly reduced, when the desirability beliefs for different age and educational groups were statistically controlled.

In the next step of our analysis, we introduced two factors, which were assumed to represent how similar different sub-samples in the response dataset are to the respondents who provided the desirability beliefs (cf. model 5, table 7). As already shown in the descriptive analysis, respondents from the western part of Germany and those who lived in communities with more than 100.000 inhabitants gave significantly more positive racial attitude answers. It was assumed that the desirability beliefs differ between the regions in Germany and accordingly to the urbanity of the respondents' home community. Thus, matched desirability beliefs are expected to have a stronger effect on the attitude responses for those subjects who are more similar to the West German and urban sample, in which these beliefs were collected. Thus, we computed interaction terms between each of both factors and the relative desirability beliefs. These terms proved in regression model 6 to be significant moderator variables for how strongly the matched desirability beliefs explained the response behavior.

- table 7 about here -

In table 8, we present racial attitude scores, predicted from the significant interaction effects in regression model 6. These attitude scores were computed for respondents who perceive incentives for positive and negative racial attitudes of a strength as we found them on average in our sample (relative desirability beliefs of -2.1 and $+2.1$). Both moderator variables had the expected effect on how strongly the desirability beliefs predicted the racial attitude answers. Accordingly, when respondents from the western part of Germany believed negative attitudes to be more desirable the attitude score was 1.6, but when positive attitudes were regarded as more favorably evaluated, this value was 5.1. This is a social desirability effect of 3.5 points on the attitude response scale. In contrast, the answers of respondents from eastern Germany are only 2.8 points sensitive to differences in the desirability beliefs, which were reported by the West German sample. Similar differences in the effect of desirability beliefs, which were obtained from the urban sample of our own study, on the responses in the ALLBUS survey were observed for subjects from bigger and smaller cities. Whereas the response behavior of subjects from cities with more than 100.000 inhabitants was affected by 3.3 points on the response scale, subjects from smaller communities were only sensitive 2.8 points to differences in desirability beliefs. We can thus conclude that the desirability beliefs had for both dimensions stronger effects on attitude answers of respondents who were more similar to those who provided the relative desirability scores.

- table 8 about here -

Summary and discussion

In this paper we analyzed the differences in the perceived desirability of positive and negative racial attitude answers. This was done for 10 items from the German General Social Survey. The motivation for this analysis was first to test differences in the incentives for socially desirable response behavior between the items. On the basis of these results, researchers can select those items which are the least susceptible to SD-bias. Our second aim was to analyze a possible heterogeneity in the strength and direction of desirability beliefs according to the respondents' demographic characteristics. Here, we are concerned about the possibility that such differences may have a systematic effect on the racial attitude reports, and that the resulting differences in SD-bias are then misinterpreted as evidence for real attitude differences between the groups of respondents.

Our results have shown *firstly* that the respondents differ considerably with respect to whether positive or negative racial attitudes are regarded to be more desirable, and how strongly a particular answer was assumed to be more favorably evaluated in society. Th pre-

dicted incentives for SD-bias in different directions partly cancel each other out on the aggregate level of the sample. Thus, the possible SD-bias was largely underestimated, when, as it was typically done in previous research, only aggregated measures for each item were utilized. Under this condition, the comparison of how susceptible the items are with respect to SD-bias was found to be biased as well. A very important result in the analysis of desirability beliefs was that these beliefs differed significantly according to the respondents' sex, age and education. These differences were with respect to the strength and direction of potential SD-bias. Females, younger and more educated respondents perceived positive racial attitude answers to be more desirable, compared with the respective complementary groups.

The second question of this paper was, whether there is evidence that the desirability beliefs indeed affects racial attitude answers. In order to avoid the possibility that desirability concerns are artificially activated, we did not collect the answers about racial attitudes in the same interview as the desirability beliefs. Instead, we used the data from the ALLBUS study in 1996, where answers to the analyzed attitude items were recorded. The item differences in desirability beliefs, as well as their distinction according to the respondents' age, gender and education were matched with the attitude answers in the ALLBUS dataset. As in other studies, we found firstly significant differences in the racial attitude reports according to the subjects' sex, age and education (cf. for example: Alba & Johnson, 2000). Secondly, the desirability measure had a strong effect on response behavior: racial attitude answers were more positive, when these responses were judged to be more desirable. After controlling for these differentiations in the perceived desirability between the same demographic groups, the strength of their association with racial attitude reports was greatly reduced. This suggests that the analyzed racial attitude scale is prone to SD-bias, that this applies to a different degree for the single items and that at least a part of the observed socio-demographic differences in the attitude answers is due to SD-effects. However, consistent with other research, after group differences in SD-bias were controlled, the socio-demographic characteristics, in particular the respondents' education, still had an effect on the attitude reports (Wagner & Zick, 1995). We can thus not support the hypothesis from Jackman (1973) that survey results indicating that more educated respondents are less prejudiced, are only due to differences in SD-bias.

Another result from our study is that the desirability beliefs which were reported by subjects from a West German, metropolitan sample, predicted racial attitude responses, and thus supposedly the strength and direction of SD-bias, to a different degree. The predictive power was significantly better for the attitude answers of those respondents from the nationwide ALLBUS sample, which were from West Germany and from more urban regions, com-

pared with the complimentary groups. This result and additional evidence lead us to the conclusion that cultural differences according to the region and the urbanity of respondents' home communities influences their desirability beliefs. Also the differences in the observed associations provides additional evidence for the predictive validity of desirability beliefs for SD-bias, this shows as well, that the distribution of desirability beliefs observed in our study cannot be simply generalized at the level of the total population. This makes it necessary to conduct a nationwide representative desirability study, which makes it possible as well, to test directly the presence and effect of regional differences in desirability beliefs.

Other studies have tested, using the same racial attitude questions as in the present investigation, whether the respondents' desirability beliefs affects their response behavior to a lower degree, when the privacy of the response situation is ensured (Stocké, 2004b). This was done with an experimental design, in which the respondents reported their desirability beliefs and their racial attitudes in the same interview. The privacy of the response situation was varied by conducting half of the interviews interviewer- and half self-administered. The results have demonstrated that subjects racial attitude answers were significantly less determined by their desirability beliefs, when the interviewer was not able to observe and potentially sanction the response behavior. These results, together with those of the present study, suggest that racial attitudes should be recorded through self-administered surveys. When this is not possible, the researcher should utilize those attitude questions, which our study has shown to be least affected by SD-bias.

References

- Alba, R. & Johnson, M. (2000). Zur Messung aktueller Einstellungsmuster gegenüber Ausländern in Deutschland. In R. Alba, P. Schmidt, & M. Wasmer (Ed.), *Deutsche und Ausländer: Freunde, Fremde oder Feinde? Empirische Befunde und theoretische Erklärungen*. Wiesbaden: Westdeutscher Verlag.
- Böltken, F. (1992). Wertorientierungen in der Bundesrepublik Deutschland im regionalen und zeitlichen Vergleich. In H. Klages, H.-J. Hippler, & W. Herbert (Ed.), *Werte und Wandel: Ergebnisse und Methoden einer Forschungstradition*. Frankfurt am Main: Campus Verlag.
- Chen, P. Y., Dai, T., Spector, P. E., & Jex, S. M. (1997). Relation Between Negative Affectivity and Positive Affectivity: Effects of Judged Desirability of Scale Items and Respondents' Social Desirability. *Journal of Personality Assessment*, 69, 183-198.
- Crott, H. W. & Roßrucker, K. (1974). Erwünschtheit von Eigenschaften in Abhängigkeit von Alter, Geschlecht und Schichtzugehörigkeit. *Zeitschrift Für Entwicklungspsychologie Und*

- Pädagogische Psychologie*, 6, 241-261.
- Crowne, D. P. & Marlowe, D. (1964). *The Approval Motive. Studies in Evaluative Dependence*. New York: John Wiley.
- Duck, R. J. & Hunsberger, B. (1999). Religious Orientation and Prejudice: The Role of Religious Proscription, Right-Wing Authoritarianism. *The International Journal for the Psychology of Religion*, 9, 157-179.
- Edwards, A. L. (1957). *The Social Desirability Variable in Personality Assessment and Research*. New York: Dryden Press.
- Epstein, J. F., Barker, P. R., & Kroutil, L. A. (2001). Mode Effects in Self-Reported Mental Health Data. *Public Opinion Quarterly*, 65, 529-549.
- Esser, H. (1991). Die Erklärung systematischer Fehler in Interviews: Befragtenverhalten als „rational choice“. In R. Wittenberg (Ed.), *Person - Situation - Institution - Kultur. Günter Büschges zum 65. Geburtstag*. Berlin: Duncker & Humblot.
- 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.
- Fuchs, D. (2000). Die demokratische Gemeinschaft in den USA und in Deutschland. In J. Gerhards (Ed.), *Die Vermessung kultureller Unterschiede. USA und Deutschland im Vergleich*. Wiesbaden: Westdeutscher Verlag.
- Gove, W. R. & Geerken, M. R. (1977). Response Bias in Surveys of Mental Health: An Empirical Investigation. *American Journal of Sociology*, 82, 1289-1317.
- Hampson, S. E., Goldberg, L. R., & John, O. P. (1987). Category-Breadth and Social-Desirability Values for 573 Personality Terms. *European Journal of Personality*, 1, 241-258.
- Herbert, W. (1991). Die Wertorientierungen der Deutschen vor der Vereinigung. Eine empirische Momentaufnahme in Ost und West. *BISS Public*, 2, 119-132.
- Huang, C., Liao, H., & Chang, S.-H. (1998). Social Desirability and the Clinical Self-Report Inventory: Methodological Reconsideration. *Journal of Clinical Psychology*, 54, 517-528.
- Jackman, M. R. (1973). Education and Prejudice or Education and Response-Set? *American Sociological Review*, 38, 327-339.
- Krysan, M. (1998). Privacy and the Expression of White Racial Attitudes: A Comparison Across Three Contexts. *Public Opinion Quarterly*, 62, 506-544.
- Krysan, M., Schuman, H., Scott, L. J., & Beatty, P. (1994). Response Rates and Response Content in Mail Versus Face-to-Face Surveys. *Public Opinion Quarterly*, 58, 381-399.

- Mecklenbräuker, S., Hager, W., & Möller, H. (1994). Erwünschtheit von 908 Eigenschaften bei Frauen und bei Männern aus der Sicht von Männern und von Frauen. In W. Hager & M. H. Hasselhorn (Eds.), *Handbuch deutschsprachiger Wortnormen*. Göttingen: Hogrefe.
- Meleddu, M. & Guicciardi, M. (1998). Self-Knowledge and Social Desirability of Personality Traits. *European Journal of Personality*, *12*, 151-168.
- Mielke, R. (1995). *Der Bildungsgrad und die Einstellung gegenüber Ausländern*. Universität Bielefeld: Bielefelder Arbeiten zur Sozialpsychologie 173.
- Mummendey, H. D., Bolten, H.-G., & Isermann-Gerke, M. (1982). Experimentelle Überprüfung des Bogus-Pipeline-Paradigmas: Einstellungen gegenüber Türken, Deutschen und Holländern. *Zeitschrift für Sozialpsychologie*, *13*, 300-311.
- Nicotera, A. M. (1996). An Assessment of the Argumentativeness Scale For Social Desirability Bias. *Communication Reports*, *9*, 23-36.
- Paulhus, D. L. & Reid, D. B. (1991). Enhancement and Denial in Socially Desirable Responding. *Journal of Personality and Social Psychology*, *60*, 307-317.
- Phillips, D. L. & Clancy, K. J. (1972). Some Effects of "Social Desirability" in Survey Studies. *American Journal of Sociology*, *77*, 921-940.
- Presser, S. & Stinson, L. (1998). Data Collection Mode and Social Desirability Bias in Self-Reported Religious Attendance. *American Sociological Review*, *63*, 137-145.
- Reinecke, J. (1991). *Interviewer- und Befragtenverhalten. Theoretische Ansätze und methodische Konzepte*. Opladen: Westdeutscher Verlag.
- Schlenker, B. R., Bonoma, T. V., Hutchinson, D., & Burns, L. (1976). The Bogus Pipeline and Stereotypes toward Blacks. *The Journal of Psychology*, *96*, 319-329.
- Sigall, H. & Page, R. (1971). Current Stereotypes: a Little Fading, a Little Faking. *Journal of Personality and Social Psychology*, *18*, 247-255.
- Singer, E. (1981). *Reference Groups and Social Evaluations*. New York: Basic Books.
- Smith, T. W. (1992). Discrepancies Between Men and Women in Reporting Number of Sexual Partners: A Summary from Four Countries. *Social Biology*, *39*, 203-211.
- Stocké, V. (2001). *Socially Desirable Response Behavior as Rational Choice: The Case of Attitudes Towards Foreigners*. Working-Paper No. 01-22: Sonderforschungsbereich 504, University of Mannheim.
- Stocké, V. (2004a). Entstehungsbedingungen von Antwortverzerrungen durch soziale Erwünschtheit. Ein Vergleich der Prognosen der Rational-Choice Theorie und des Modells der Frame-Selektion. *Zeitschrift für Soziologie*, *33*, 303-320.
- Stocké, V. (2004b). *The Interdependence of Determinants for the Strength and Direction of*

- Social Desirability Bias in Racial Attitude Surveys*. Working Paper No. 04-14: Sonderforschungsbereich 504, University of Mannheim.
- Stocké, V. & Hunkler, C. (2004). Die angemessene Erfassung der Stärke und Richtung von Anreizen durch soziale Erwünschtheit. *ZA-Information*, 54, 53-88.
- Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The Psychology of Survey Response*. Cambridge: Cambridge University Press.
- Wagner, U. & Zick, A. (1995). The Relation of Formal Education to Ethnic Prejudice: Its Reliability, Validity and Explanation. *European Journal of Social Psychology*, 25, 41-56.
- Wyner, G. A. (1980). Response Errors in Self-Reported Number of Arrests. *Sociological Methods & Research*, 9, 161-177.

Footnotes

- ¹ Whereas in this theory, the SD-bias is assumed to be motivated by the subjects' impression management motive, in other theoretical approaches the respondents' motivation for 'self deception' is regarded as important as well. Here, independent from the demands perceived from the side of others, subjects unconsciously bias their reports in the direction of social desirability, in order to preserve a positive self-image (cf. for a discussion of the difference between 'self deception' and 'other deception': Paulhus & Reid, 1991).
- ² The dataset is available from the German Central Archive, study number ZA 2800. The items 1 to 4 have been used in the years 1980, 1984, 1988, 1990, 1994 and 2000 as well.
- ³ These desirability scores are the predicted values for the socio-demographic groups, when all other variables included in the regression equation are fixed on their reference categories.

Table 1. Respondents' demographic characteristics

	Trait Desirability Study Percent (N)	ALLBUS 1996 Percent (N)
Sex		
- Male	47.9 (181)	48.9 (1578)
- Female	52.1 (197)	51.1 (1651)
Education		
- Compulsory education	39.4 (149)	46.7 (1507)
- Secondary school certificate	33.1 (125)	29.5 (952)
- High school certificate	17.2 (65)	11.1 (395)
- University degree	10.3 (39)	12.7 (411)
Mean age	43.7	46.6
All	100.0 (378)	100.0 (3229)

Table 2. Absolute and relative desirability beliefs aggregated on the item level

	Desirability of positive attitude answers *)	Desirability of negative attitude answers *)	Relative desirability **)
	Mean (Std.)	Mean (Std.)	Mean (Std.)
Item 1 “Foreigners in Germany should adapt their lifestyle more to that of Germans.”	.41 (1.4)	.52 (1.9)	-.11 (2.6)
Item 2 ^R „Foreigners in Germany should be entitled to the same social welfare and other social security benefits as Germans.“	.08 (1.6)	.05 (1.6)	.03 (3.0)
Item 3 “Foreigners in Germany commit more criminal offenses than Germans.”	.37 (1.3)	-.24 (1.7)	.61 (2.7)
Item 4 “The presence of foreigners in Germany causes problems in the housing market.”	.41 (1.2)	-.29 (1.4)	.70 (2.1)
Item 5 “Foreigners in Germany are a burden to the social security system.”	.57 (1.4)	-.47 (1.9)	1.03 (3.0)
Item 6 “One should forbid any political activities of foreigners in Germany.”	.59 (1.4)	-.50 (1.7)	1.08 (2.7)
Item 7 “Foreigners in Germany should marry among their own compatriots.”	.81 (1.4)	-.45 (1.6)	1.25 (2.4)
Item 8 “In case of increasing unemployment, foreigners should be sent back to their home countries.”	.60 (1.4)	-.70 (1.7)	1.28 (2.6)
Item 9 “Because of so many foreigners in Germany, one feels a stranger in one’s own country.”	.95 (1.3)	-.64 (1.5)	1.59 (2.3)
Item 10 ^R . „Foreigners in Germany are an enrichment for our culture.“	.92 (1.5)	-.69 (1.4)	1.61 (2.4)
Total	.57 (1.4)	-.34 (1.7)	.91 (2.6)
Number of cases	N=378		

Note: * Response scale ranges from -4 (attitude undesirable) to +4 (attitude desirable); ** Scale ranges between -8 (negative attitude more desirable) and +8 (positive attitude more desirable). For items marked with ‘R’ the original coding of responses was recoded in a way consistent with the other questions.

Table 3. Heterogeneity of relative desirability beliefs about positive and negative racial attitudes

	Positive attitudes more desirable ^{*)}		No differences	Negative attitudes more desirable ^{**)}		Total incentives ^{***)}
	%	Mean (Std)		%	%	
Item 1	34.4	2.6 (1.7)	25.9	39.7	-2.5 (1.6)	1.9 (1.8)
Item 2	39.7	2.7 (2.0)	19.8	40.5	-2.6 (1.9)	2.1 (2.0)
Item 3	49.5	2.7 (1.7)	20.9	29.6	-2.5 (1.7)	2.1 (1.8)
Item 4	50.3	2.3 (1.6)	28.3	21.4	-2.1 (1.2)	1.6 (1.6)
Item 5	58.0	3.0 (1.7)	16.7	25.4	-2.8 (1.8)	2.5 (1.9)
Item 6	60.3	2.7 (1.7)	16.7	23.0	-2.4 (1.8)	2.2 (1.9)
Item 7	55.3	2.9 (1.7)	26.7	18.0	-1.9 (1.3)	1.9 (1.9)
Item 8	62.2	2.8 (1.7)	18.3	19.6	-2.4 (1.5)	2.2 (1.8)
Item 9	62.2	3.0 (1.7)	23.3	14.6	-1.8 (1.3)	2.1 (1.9)
Item 10	64.8	3.0 (1.7)	20.6	14.6	-2.2 (1.3)	2.3 (1.9)
Total	53.7	2.8 (1.7)	21.7	24.6	-2.4 (1.6)	2.1 (1.9)
N	378					

Note: ^{*)} Relative desirability values ranges from greater than zero to +8 (positive attitudes more desirable). ^{**)} Values range from smaller than zero to -8 (negative attitudes more desirable). ^{***)} Total incentives are the average of respondents' absolute value of their relative desirability scores. Subjects with no desirability differences are included in this measure. This variable ranges from 0 (no incentives) to 8 (strong incentives).

Table 4. Differences in incentives for SD-bias between attitude items, taking the individual level heterogeneity in the direction of incentives into account (results of two-sided t-tests; N=378)

	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9	Item 10
Item 1	--									
Item 2	t=2.1 p ≤ 0.05	--								
Item 3	t=2.0 p ≤ 0.05	t=0.3 p > 0.05	--							
Item 4	t=2.7 p ≤ 0.05	t=5.2 p ≤ 0.05	t=5.1 p ≤ 0.05	--						
Item 5	t=4.8 p ≤ 0.05	t=3.0 p ≤ 0.05	t=3.6 p ≤ 0.05	t=8.5 p ≤ 0.05	--					
Item 6	t=2.6 p ≤ 0.05	t=0.4 p > 0.05	t=0.7 p > 0.05	t=6.3 p ≤ 0.05	t=2.5 p ≤ 0.05	--				
Item 7	t=0.4 p > 0.05	t=1.8 p > 0.05	t=1.5 p > 0.05	t=3.6 p ≤ 0.05	t=4.5 p ≤ 0.05	t=2.5 p ≤ 0.05	--			
Item 8	t=2.9 p ≤ 0.05	t=0.6 p > 0.05	t=1.0 p > 0.05	t=6.7 p ≤ 0.05	t=2.2 p ≤ 0.05	t=0.3 p > 0.05	t=3.0 p ≤ 0.05	--		
Item 9	t=1.9 p > 0.05	t=0.3 p > 0.05	t=0.0 p > 0.05	t=5.4 p ≤ 0.05	t=3.2 p ≤ 0.05	t=0.7 p > 0.05	t=1.8 p > 0.05	t=1.0 p > 0.05	--	
Item 10	t=3.1 p ≤ 0.05	t=1.0 p > 0.05	t=1.2 p > 0.05	t=6.5 p ≤ 0.05	t=1.7 p > 0.05	t=0.7 p > 0.05	t=3.1 p ≤ 0.05	t=0.4 p > 0.05	t=1.4 p > 0.05	--

Table 5. Mean response behavior in the ALLBUS study according to region and size of respondents' home community

	Region in Germany		Inhabitants of home community		Total
	East	West	less than 100.000	more than 100.000	
	Mean	Mean	Mean	Mean	
Item 1	3.0	3.1	3.0	3.3	3.1
Item 2 ^R	4.3	4.4	4.3	4.6	4.4
Item 3	3.5	4.1	3.8	4.2	3.9
Item 4	4.0	3.7	3.7	3.9	3.8
Item 5	3.4	4.1	3.7	4.1	3.8
Item 6	4.1	4.2	4.1	4.5	4.2
Item 7	4.7	5.3	4.9	5.6	5.1
Item 8	3.9	4.7	4.3	4.8	4.4
Item 9	4.7	4.8	4.7	5.0	4.8
Item 10 ^R	3.7	3.9	3.6	4.2	3.8
Total	3.9	4.2	4.0	4.4	4.1
Number of cases	1092	2137	2211	1018	3229

Note: For items marked with 'R' the original coding of the responses was recoded in a way that high values express a positive racial attitude. Response scale ranges in all cases between 1 (negative attitudes) and 7 (positive attitudes). Data is based on the 1996 German General Social Survey.

Table 6. Differences in relative desirability beliefs between items and respondents' social characteristics (multi-level model fixed effects with random effects for respondents)

	Model 1 B (t-Value)		Model 2 B (t-Value)
<i>Fixed effects</i>		<i>Fixed effects</i>	
Item differences ^{a)}		Item differences ^{a)}	
- Item 2	.13 (0.8)	- Item 2	.13 (0.8)
- Item 3	.72 (4.5) *	- Item 3	.72 (4.5) *
- Item 4	.81 (5.1) *	- Item 4	.81 (5.1) *
- Item 5	1.14 (7.2) *	- Item 5	1.14 (7.2) *
- Item 6	1.19 (7.5) *	- Item 6	1.19 (7.5) *
- Item 7	1.36 (8.6) *	- Item 7	1.36 (8.6) *
- Item 8	1.39 (8.7) *	- Item 8	1.39 (8.7) *
- Item 9	1.70 (10.7) *	- Item 9	1.70 (10.7) *
- Item 10	1.72 (10.8) *	- Item 10	1.72 (10.8) *
<i>Group differences</i>		<i>Group differences</i>	
Age ^{b)}		Age ^{b)}	
- 18-32 years	.51 (2.3) *	} 18-50 years	.46 (2.7) *
- 33-50 years	.49 (2.4) *		
Male ^{c)}	-.32 (2.1) *	Male ^{c)}	-.32 (2.1) *
Education ^{d)}		Education ^{e)}	
- Compulsory education	-.45 (1.6)	} Compulsory education & Secondary school certificate	-.67 (3.7) *
- Secondary school certificate	-.58 (2.1) *		
- High school certificate	.22 (0.7)		--
Constant	.04 (0.1)	Constant	.22 (0.8)
σ^2 – level 1	2.19	σ^2 – level 1	2.19
σ^2 – level 2	1.32	σ^2 – level 2	1.31
R ² – within	0.07	R ² – within	0.07
R ² – between	0.08	R ² – between	0.08
Wald χ^2	286.9 (df 14) *	Wald χ^2	286.0 (df 12) *

Note: Significance: * $p \leq 0.05$; Omitted categories: ^{a)} item 1, ^{b)} over 50 years, ^{c)} female, ^{d)} university degree, ^{e)} high school certificate & university degree.

Table 7. Effect of desirability beliefs on racial attitude answers in the ALLBUS survey in 1996 (multi-level fixed effects with random effects for respondents controlled)

	Model 3 B (t-value)	Model 4 B (t-value)	Model 5 B (t-value)	Model 6 B (t-value)
<i>Fixed effects</i>				
<i>Control variables</i>				
<i>Item</i> ^{a)}				
- Item 2	1.27 (30.5)*	1.15 (22.6)*	1.16 (22.8)*	1.16 (22.8)*
- Item 3	.83 (20.0)*	.21 (1.3)	.26 (1.6)	.25 (1.6)
- Item 4	.71 (17.2)*	.02 (0.1)	.07 (0.4)	.06 (0.4)
- Item 5	.74 (17.9)*	-.25 (1.0)	-.17 (0.7)	-.18 (0.7)
- Item 6	1.10 (26.6)*	.07 (0.3)	.15 (0.6)	.14 (0.5)
- Item 7	2.05 (49.3)*	.87 (2.9)*	.96 (3.2)*	.95 (3.1)*
- Item 8	1.34 (32.2)*	.13 (0.4)	.22 (0.7)	.22 (0.7)
- Item 9	1.69 (40.5)*	.21 (0.6)	.32 (0.9)	.31 (0.8)
- Item 10	.75 (18.0)*	-.74 (1.9)	-.63 (1.7)	-.64 (1.7)
<i>Age</i> ^{b)}				
- 18-32 years	.57 (10.1)*	.22 (2.0)*	.22 (2.0)*	.21 (2.0)
- 33-50 years	.43 (8.3)*	.07 (0.7)	.09 (0.8)	.08 (0.8)
<i>Male</i> ^{c)}				
	.09 (2.0)*	.36 (4.4)*	.34 (4.2)*	.34 (4.2)*
<i>Education</i> ^{d)}				
- Compulsory education	-1.16 (17.3)*	-.57 (3.5)*	-.59 (3.6)*	-.58 (3.5)*
- Secondary school certificate	-.75 (10.6)*	-.17 (1.1)	-.16 (1.0)	-.14 (0.9)
- High school certificate	-.25 (2.9)*	-.25 (2.9)*	-.28 (3.2)*	-.28 (3.3)*
<i>Explanatory variables</i>				
Relative trait <u>desirability</u> (matched index values)	--	.87 (3.9)*	.80 (3.7)*	.67 (3.0)*
Region (west) ^{e)}	--	--	.28 (6.2)*	.15 (2.9)*
Community size (> 100.000) ^{f)}	--	--	.22 (5.0)*	.13 (2.3)*
Desirability • Region	--	--	--	.16 (4.8)*
Desirability • Community size	--	--	--	.11 (3.4)*
Constant	3.56 (47.8)*	3.33 (34.9)*	3.07 (30.9)*	3.18 (31.5)*
σ^2 – level 1	1.67	1.67	1.67	1.67
σ^2 – level 2	1.06	1.05	1.04	1.04
R ² – within	0.11	0.11	0.11	0.11
R ² – between	0.18	0.19	0.20	0.21
Wald χ^2	4167.2 (df. 15)*	4185.6 (df. 16)*	4273.1 (df. 18)*	4317.2 (df. 20)*

Note: Significance: * $p \leq 0.05$; omitted categories: ^{a)} item 1, ^{b)} over 50 years, ^{c)} female, ^{d)} university degree, ^{e)} east, ^{f)} < 100.000.

Table 8. Differences in the effect of relative desirability beliefs on the racial attitude reports from the ALLBUS study in 1996 according to the region and urbanity of the respondents' home community (predicted values from regression model 6)

	<i>Respondents' relative desirability beliefs</i>		Predictive power of desirability beliefs for attitude reports
	negative attitudes more desirable	positive attitudes more desirable	
<i>Region</i>			
West Germany	1.6	5.1	3.5
East Germany	1.8	4.6	2.8
<i>Community size</i>			
More than 100.000 inhabitants	1.7	4.9	3.3
Less than 100.000 inhabitants	1.8	4.6	2.8

Note: The predicted values are computed for relative desirability values of -2.1 (negative attitudes more desirable) and $+2.1$ (positive attitudes more desirable). These values represents the mean total incentives from social desirability we found in section 5.2, table 4. In the case of each of the moderator variables the respective other moderator variables and all control variables were fixed at their reference category.

SONDERFORSCHUNGSBereich 504 WORKING PAPER SERIES

Nr.	Author	Title
04-43	Fabian Bornhorst Andrea Ichino Oliver Kirchkamp Karl H. Schlag Eyal Winter	How do People Play a Repeated Trust Game? Experimental Evidence
04-42	Martin Hellwig	Optimal Income Taxation, Public-Goods Provision and Public-Sector Pricing: A Contribution to the Foundations of Public Economics
04-41	Thomas Gschwend	Comparative Politics of Strategic Voting: A Hierarchy of Electoral Systems
04-40	Ron Johnston Thomas Gschwend Charles Pattie	On Estimates of Split-Ticket Voting: EI and EMax
04-39	Volker Stocké	Determinants and Consequences of Survey Respondents' Social Desirability Beliefs about Racial Attitudes
04-38	Siegfried K. Berninghaus Marion Ott Bodo Vogt	Restricting the benefit flow from neighbors: Experiments on network formation
04-37	Christopher Koch	Behavioral Economics und die Unabhängigkeit des Wirtschaftsprüfers - Ein Forschungsüberblick
04-36	Christopher Koch	Behavioral Economics und das Entscheidungsverhalten des Wirtschaftsprüfers - Ein Forschungsüberblick
04-35	Christina Reifschneider	Behavioral Law and Economics: Überlegungen zu den Konsequenzen moderner Rationalitätskonzepte für die Gestaltung informationellen Kapitalmarktrechts
04-34	Siegfried K. Berninghaus Karl-Martin Ehrhart Marion Ott Bodo Vogt	Searching for "Stars" - Recent Experimental Results on Network Formation -