Are Racial Stereotypes Really Fading?
The Princeton Trilogy Revisited

Patricia G. Devine • University of Wisconsin–Madison
Andrew J. Elliott • University of Rochester

In this article, the authors identify three methodological shortcomings of the classic Princeton trilogy studies: (a) ambiguity of the instructions given to respondents, (b) no assessment of respondents’ level of prejudice, and (c) use of an outdated list of adjectives. These shortcomings are addressed in the authors’ assessment of the stereotype and personal beliefs of a sample of University of Wisconsin students. In contrast to the commonly espoused fading stereotype proposition, data suggest that there exists a consistent and negative contemporary stereotype of Blacks. Comparing the data from the Princeton trilogy studies with those of the present study, the authors conclude that the Princeton trilogy studies actually measured the respondents’ personal beliefs, not (as typically assumed) their knowledge of the Black stereotype. Consistent with Devine’s model, high- and low-prejudiced individuals did not differ in their knowledge of the stereotype of Blacks but diverged sharply in their endorsement of the stereotype.

During the past 50 years, there have been dramatic shifts in the social and political climate in America that have made racial discrimination illegal and overt expressions of racial prejudice socially taboo. In the wake of these macro-level changes, many social scientists have been interested in the extent to which individual-level social stereotype change has followed suit. These considerations are important in intergroup relations research, as stereotype change is commonly considered a prerequisite to prejudice reduction and more amicable intergroup interactions. Based on data from the adjective checklist procedure, the classic stereotype assessment technique developed by Katz and Braly (1933), many researchers have concluded that individual racial stereotypes in America have generally faded over the years. In the present research, we offer a critique of Katz and Braly’s procedure, guided by the question, Are individual racial stereotypes in America really fading?

In Katz and Braly’s (1933) adjective checklist procedure, respondents are provided with a list of 84 trait adjectives and instructed “to read through the list of words and select those which seem to you typical of (target group)” (p. 283). Respondents are encouraged to select as many adjectives as necessary to capture the target group adequately and are encouraged to add their own adjectives as needed. Respondents are then asked to identify the five words in the list that seem to them most typical of members of the target group. The content of the stereotype of a group is defined as the set of adjectives that respondents most frequently assign to the group.
In their seminal study, Katz and Braly (1933) found a high level of consistency in the adjectives respondents associated with the Black stereotype. In addition, the majority of the adjectives selected were fairly negative (e.g., superstitious, lazy, ignorant). Several studies have replicated Katz and Braly’s procedure in efforts to examine the amount of stability or change in stereotypes. Two of the studies (Gilbert, 1951; Karlins, Coffman, & Walters, 1969), together with Katz and Braly’s study, have been referred to as the Princeton trilogy, because they assessed stereotypes held by three generations of Princeton students. The two later studies in the Princeton trilogy used the same procedure and the same set of adjectives employed in Katz and Braly’s initial investigation. Many years later, Dovidio and Gaertner (1986), although not studying Princeton students, reported another investigation of the Black stereotype that followed Katz and Braly’s procedures exactly.1

In summarizing the pattern of findings for the Black stereotype, Dovidio and Gaertner (1991) stated that “adjective checklist studies, in which respondents are asked to select traits that are the most typical of particular racial or ethnic categories, indicate that negative stereotypes are consistently fading” (p. 202; see also Gilbert, 1951, and Karlins et al., 1969). That is, over time, respondents have selected a different set of traits to represent the Black stereotype, they have displayed less consistency in the adjectives selected, and the traits they have chosen have been less negatively valued. These findings could be interpreted as very encouraging for intergroup relations. They suggest that stereotype change is possible and that White Americans are becoming more positive in their perceptions of Blacks. Changes in the social and political climate, then, do appear to be associated with individual-level changes. However, we believe that there is some ambiguity in exactly what has changed over time. In reviewing this literature, we found that Katz and Braly’s (1933) adjective checklist procedure as it has been traditionally implemented suffers from several shortcomings that make conclusions about stereotype persistence or change tenuous at best. In what follows, we outline these shortcomings and report a study designed to address each of the shortcomings in an effort to document the nature of Whites’ representations of Black Americans as a social group.

The methodological shortcomings of the adjective checklist procedure fall into three general categories: (a) ambiguity of the instructions given to respondents, (b) no assessment of respondents’ level of racial prejudice, and (c) use of an outdated list of adjectives. In addition to these methodological shortcomings, there has been a conspicuous absence of attention to theoretical and conceptual considerations in employing Katz and Braly’s (1933) procedure and in stereotype assessment research more generally (Ashmore & Del Boca, 1981). In the following section, we briefly review Devine’s (1989) model of automatic and controlled processes in stereotyping and prejudice, which highlights a conceptual distinction in need of consideration in the stereotype assessment literature.

The Distinction Between Stereotypes and Personal Beliefs

Devine (1989), in her model of automatic and controlled processes in prejudice, argued that stereotypes and personal beliefs are conceptually distinct cognitive structures and that each structure represents only part of an individual’s entire knowledge base of a particular group. Devine suggested that a stereotype is a well-learned set of associations that link a set of characteristics with a group label (see also Dovidio, Evans, & Tyler, 1986). Through the socialization process, all individuals learn a variety of cultural stereotypes that become part of their associative network (Ehrlich, 1973). Although everyone possesses knowledge of numerous stereotypes, not all possess personal beliefs that are congruent with these stereotypes. Devine argued that beliefs are propositions that are endorsed and accepted as true. Therefore, whereas most White Americans possess knowledge of the Black stereotype (i.e., the stereotype exists within their associative network), only a subset of these individuals actually endorse the stereotype.

---

1Dovidio and Gaertner (1986) reported their stereotype assessment data in a chapter and only provided a comparison of the traits most frequently selected by their respondents with those most frequently selected in the Princeton trilogy studies. Details about the number of respondents that participated in the study, the percentage of respondents who refused to complete the task, and so forth are not provided in their chapter.
and believe it is veracious (for a comparable argument in the attitude theory literature, see Pratkanis, 1989; Zanna & Rempel, 1988). For high-prejudiced individuals, their knowledge of the Black stereotype and their personal beliefs about Blacks are highly congruent, whereas low-prejudiced individuals have rejected the stereotype and adopted a distinctly different set of personal beliefs about Blacks. Importantly, however, rejection of the Black stereotype by low-prejudiced individuals does not immediately eradicate the stereotype from their associative network. On the contrary, the stereotype often remains a well-organized, frequently activated knowledge structure (see Devine, 1989, Study 2; Jamieson & Zanna, 1989; Klinger & Beall, 1992; Kruglanski & Freund, 1983; Sherman & Gorkin, 1980).

Devine (1989, Study 1) provided direct empirical evidence that both high- and low-prejudiced respondents possess equivalent knowledge of the Black stereotype. In a free-response task, respondents were instructed to list the components of the Black stereotype. Respondents were informed that the researchers were not interested in their personal beliefs about Blacks but only their knowledge of the Black stereotype. Comparisons of the trait ascriptions generated by high- and low-prejudiced respondents failed to yield any significant differences. Indirect support for the proposition that high- and low-prejudiced individuals possess different personal beliefs toward Blacks was obtained in a subsequent study (Devine, 1989, Study 3). Under conditions emphasizing anonymity, respondents were instructed to list their thoughts about Black Americans as a social group. Analysis of these free-response data yielded clear differences between high- and low-prejudiced respondents: Those high in prejudice listed more negative traits than did those low in prejudice. In general, low-prejudiced respondents’ thoughts reflected themes of equality and negation of the stereotype. Together, these studies support the conceptual distinction between stereotypes and personal beliefs.

**Shortcomings of the Princeton Trilogy Studies**

The stereotype/personal-belief distinction is especially relevant to the first methodological shortcoming of the Princeton trilogy—the ambiguity of instructions. Both Gilbert (1951) and Karlin et al. (1969) reported that many of their respondents found Katz and Braly’s (1933) task confusing. From their written comments about the task (which respondents were encouraged to provide), it appears that some respondents thought the instructions asked for them to list their knowledge of the stereotype, whereas others thought they were being instructed to list their personal beliefs. Given the ambiguity of the instructions, it is difficult to know precisely what it is that respondents reported—their knowledge of the stereotype or their personal beliefs. Clearly, before claims about stereotype change can be made from the Princeton trilogy data, it is important to ascertain whether respondents indeed reported their knowledge of the stereotype. An equally plausible possibility, from Devine’s (1989) perspective, is that the stereotype has remained stable through the years (in consistency and valence, not necessarily in specific content), whereas personal beliefs have undergone a revision. In the present research, we sought to eliminate ambiguity by providing respondents with clear instructions concerning whether they should list their knowledge of the Black stereotype or their personal beliefs about Blacks. In fact, all respondents completed both stereotype and personal beliefs measures so that respondents’ knowledge of the stereotype could be compared with their personal beliefs.

Our second methodological criticism of the Princeton trilogy is also driven by a consideration of the stereotype/personal-belief distinction. In none of the studies in the Princeton trilogy was there any independent assessment of respondents’ level of prejudice toward the target group. From Devine’s (1989) perspective, the assessment of prejudice is of fundamental importance, given that high- and low-prejudiced individuals are equally knowledgeable of the Black stereotype but diverge in their endorsement of the stereotype. Consideration of prejudice level, therefore, can afford a much more precise analysis of respondents’ representation of the target group. Karlin et al. (1969) recognized the value of employing an independent indicator of prejudice in research on stereotypes and attitudes, and actually called for future research to attend more closely to differences between high- and low-prejudiced individuals. In the present research, we heeded the call of Karlin et al. by assessing prejudice with McConahay, Hardee, and Batts’ (1981) Modern Racism Scale (MRS).

Our third criticism of the Princeton trilogy stud-
of their respondents' task confusion about the task to provide), thought the in-knowledge ought they were ul beliefs. Given it is difficult to respondents re-stereotype or their aims about ste-n the Princeton certain whether knowledge of ble possibility, is that the ste-ghe the years (in asserity in spectacles have un-research, we providing re-nts concerning wldge of the l beliefs about ompleted both essures so that stereotype could eliefs.

riticism of the consideration distinction. In on trilogy was of respondents’ st group. From assessment of ice, given that als are equally ype but diverge ype. Consider-
nes, can afford a ndents' repres et al. (1969) an independent stereotypes and ure research to ween high- and esent research, l. by assessing rce, and Batts's S).

ies is that they all employed the same list of 84 adjectives used in Katz and Braly’s (1933) initial assessment. Karlins et al. (1969) expressed concern over the adequacy of the 1933 list for capturing the then current stereotype: “We suspect that a comprehensive list for today would have to include several terms not found in the 1933 version, terms which might have been widely subscribed to by our subjects” (p. 14). Although respondents were encouraged to add traits to the list, Karlins et al. argued that the constraints of a provided list are strong and that typically few respondents depart from the list. Devine’s (1989, Study 1) stereotype assessment study validated the concerns of Karlins et al., as many of the characteristics most frequently generated by respondents to capture the stereotype of Blacks (e.g., poor, criminal, athletic) were not included on the original Katz and Braly list.

Karlins et al. (1969) argued that the reason they used the outdated list of adjectives (and the same set of instructions) was to ensure comparability across the studies (their study being the third in the trilogy). They wanted to ensure that any differences found in the pattern of trait selection were not due to changes in the task. However, change in a stereotype may involve the addition, as well as deletion, of traits from the original characterization of Blacks suggested in Katz and Braly's (1933) study. Therefore, in the present research, we expanded the list of adjectives provided to respondents to include the traits frequently listed by Devine’s respondents (1989, Study 1) that were not included in the original Katz and Braly list.

In summary, the primary goal of the present research was to assess the contemporary stereotype of Blacks and the personal beliefs about Blacks among a sample of University of Wisconsin undergraduates, using a revised Katz and Braly procedure, one that addressed the previously discussed methodological shortcomings. In so doing, we hoped to provide evidence attesting to the veracity of our methodological critique and to offer a reevaluation of the fading stereotype proposition currently being espoused on the basis of the Princeton trilogy data. Given the presumed implications of stereotype change for the nature of intergroup relations, it is important to develop a clear analysis of exactly what has and has not changed over time (and for whom). In addition, the present study allowed us to replicate Devine’s (1989) stereotype assessment study and, more important, to directly test the proposition that high- and low-prejudiced respondents would report differential personal beliefs, despite their possessing equivalent knowledge of the stereotype.

Method

Respondents and Design

Respondents were 147 White University of Wisconsin-Madison undergraduates who received extra credit in their introductory psychology class in return for their participation. All respondents completed both the stereotype and the personal beliefs assessments. Half the respondents completed the stereotype assessment first and then the personal beliefs assessment, whereas the other half completed the tasks in the reverse order. Respondents were randomly assigned to one of the two orders. A nonreactive measure of anti-Black attitudes was administered at the end of the session, and respondents were assigned to a high-prejudice (n = 74) or low-prejudice (n = 73) group based on a median split of the scores obtained in the sample.

Procedure

Respondents first read and signed a consent form stating that the purpose of the research was to understand people’s knowledge of and reactions to various social groups and that in the present study, the focus was on Blacks. They were told that all their responses would be kept anonymous and that their participation was voluntary. After providing respondents with a booklet containing the materials and instructions, the experimenter left the room. The instructions asked respondents not to write their names or any other identifying information in the booklet. Respondents were informed that at the completion of the study, each respondent would

---

2 We also conducted a pilot study in which we gave 50 respondents the Katz and Braly (1933) list of adjectives and very clear instructions to consider whether the list was adequate to capture the contemporary cultural stereotype of Blacks. Respondents were specifically instructed to think about the cultural stereotype and not their personal beliefs. They were further instructed to add any traits they thought would be necessary to capture the cultural stereotype. The types of traits added to the original list of traits overlapped almost completely with the traits frequently listed by respondents in Devine’s (1989) free-response stereotype assessment task (e.g., athletic, poor, rhythmic, uneducated, criminal, low in intelligence).
place the booklet into an unmarked envelope that would be deposited in a drop box. This procedure was designed to assure respondents that responses could not be associated with any particular individual. Respondents then completed the stereotype and personal beliefs assessments (a full description is given below). Finally, an independent assessment of respondents’ prejudice level was obtained. Upon completion of their booklet, respondents were debriefed, presented with an extra-credit card, and dismissed.

Measures

**Stereotype assessment.** The stereotype assessment instructions were designed to eliminate the ambiguity present in Katz and Braly’s (1933) instructions (cf. Gilbert, 1951; Karlins et al., 1969) by explicitly stating that we were interested in respondents’ knowledge of the cultural stereotype of Blacks. The instructions read as follows:

On the back of this page you will find a list of adjectives and on the page next to that a series of blank lines. Please read through the list carefully and identify those adjectives that make up the cultural stereotype of Blacks. Note, these characteristics may or may not reflect your personal beliefs. So, select those adjectives that you know to be part of the cultural stereotype whether or not you believe the stereotype to be true. Please list them in the blanks. If you do not find all the necessary adjectives in the list (i.e., the list is incomplete), you may add any other information that you think is necessary to represent the cultural stereotype. Use as many or as few blanks as you need.

After respondents selected adjectives and turned the page, they read the following instruction:

What we would like you to do now is to go back to the set of adjectives you selected and listed as making up the cultural stereotype. Please mark with a star the five words that seem to you to be the most typical of Blacks according to the cultural stereotype.

**Personal beliefs assessment.** The personal beliefs instructions were designed to make it clear that we were interested in respondents’ personal beliefs about Blacks. Again, our concern was to be unambiguous about what we were asking respondents to report. The instructions read as follows:

On the back of this page you will find a list of adjectives and on the page next to that a series of blank lines. Please read through the list carefully and identify those adjectives that you personally believe characterize Blacks. Please list them in the blanks. If you do not find all the necessary adjectives (i.e., the list is incomplete), you may add any other information you think is necessary to represent your personal beliefs about Blacks. Use as many or as few blanks as you need.

After respondents completed the personal beliefs assessment, they were asked to go back to the list they had generated and mark with a star the five words that seemed most typical of Blacks according to their personal beliefs.

**Adjective List.** Respondents were presented with a list of 93 adjectives. This list included the 84 original Katz and Braly (1933) adjectives as well as 9 characteristics that were strongly suggested to be part of the cultural stereotype of Blacks in a free-response task (see Devine, 1989, Study 1). The characteristics added were athletic, criminal, hostile, low in intelligence, poor, rhythmic, sexually perverse, uneducated, and violent. The 93 characteristics were randomly ordered and presented on the page in three columns.

**Prejudice Assessment.** The 7-item version of the Modern Racism Scale (MRS) was employed to assess respondents’ level of prejudice toward Blacks. This assessment device was designed to be a nonreactive measure of anti-Black attitudes and has been useful in predicting a variety of responses, including voting patterns (Sears & Kinder, 1971) and reactions to busing (Sears & McConahay, 1973). Previous research has attested to the reliability and validity of the scale (see McConahay, 1986), and the scale proved reliable in the present study (Cronbach’s alpha = .85).

The MRS items were embedded in a series of other social and political questions, and respondents indicated their agreement with each of the items (e.g., “It is easy to understand the anger of Black people in America”) on a 9-point scale that ranged from −4 (Disagree strongly) to +4 (Agree strongly). Although the possible range of total MRS scores was −28 (low prejudiced) to +28 (high prejudiced), the observed range of scores in the present sample was −28 to +9. It is common in college samples for the distribution of MRS scores to be positively skewed, and we have observed this pattern in University of Wisconsin samples during the last several years. However, despite the restriction in range in MRS scores, the scale has been sensitive to even subtle differences between
those who score extreme in the low end of the continuum and those who score in the middle to high ranges of the scale (e.g., Zuwerink, Monteith, Devine, & Cook, 1996).

Results

Preliminary Analyses

Preliminary analyses tested for effects of the order of stereotype and personal beliefs assessments. No order effects were found in any of these analyses; thus all results reported below are collapsed across the order variable.

Knowledge of Cultural Stereotype

In comparing the content of the Black stereotype across the Princeton trilogy studies, researchers have commonly adopted the frequency of trait selections in the five-starred-trait task as their unit of analysis. Table 4.1 presents these data for the Princeton trilogy studies, the 1982 Dovidio and Gaertner investigation (cited in Dovidio &

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Superstitious</td>
<td>84</td>
<td>41</td>
<td>13</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Lazy</td>
<td>75</td>
<td>31</td>
<td>26</td>
<td>12</td>
<td>45</td>
</tr>
<tr>
<td>Happy-go-lucky</td>
<td>38</td>
<td>17</td>
<td>27</td>
<td>15</td>
<td>3</td>
</tr>
<tr>
<td>Ignorant</td>
<td>38</td>
<td>17</td>
<td>11</td>
<td>10</td>
<td>14</td>
</tr>
<tr>
<td>Musical</td>
<td>26</td>
<td>33</td>
<td>47</td>
<td>29</td>
<td>11</td>
</tr>
<tr>
<td>Ostenisious</td>
<td>26</td>
<td>11</td>
<td>25</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Very religious</td>
<td>24</td>
<td>17</td>
<td>8</td>
<td>23</td>
<td>11</td>
</tr>
<tr>
<td>Stupid</td>
<td>22</td>
<td>10</td>
<td>4</td>
<td>1</td>
<td>11</td>
</tr>
<tr>
<td>Physically dirty</td>
<td>17</td>
<td></td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Naive</td>
<td>14</td>
<td></td>
<td>4</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Slovenly</td>
<td>13</td>
<td></td>
<td>5</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Unreliable</td>
<td>12</td>
<td></td>
<td>6</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Pleasure lovinga</td>
<td>–</td>
<td>19</td>
<td>26</td>
<td>20</td>
<td>1</td>
</tr>
<tr>
<td>Sensitiveb</td>
<td>–</td>
<td>17</td>
<td>13</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>Gregariousc</td>
<td>–</td>
<td></td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Talkativeb</td>
<td>–</td>
<td></td>
<td>14</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Imitatived</td>
<td>–</td>
<td></td>
<td>13</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>Aggressef</td>
<td>–</td>
<td></td>
<td>–</td>
<td>19</td>
<td>5</td>
</tr>
<tr>
<td>Materialisticg</td>
<td>–</td>
<td></td>
<td>–</td>
<td>19</td>
<td>0</td>
</tr>
<tr>
<td>Loyal to familyh</td>
<td>–</td>
<td></td>
<td>–</td>
<td>39</td>
<td>11</td>
</tr>
<tr>
<td>Arroganti</td>
<td>–</td>
<td></td>
<td>–</td>
<td>14</td>
<td>3</td>
</tr>
<tr>
<td>Ambitiousj</td>
<td>–</td>
<td></td>
<td>–</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>Tradition lovingk</td>
<td>–</td>
<td></td>
<td>–</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>Athleticl</td>
<td>–</td>
<td></td>
<td>–</td>
<td>74</td>
<td></td>
</tr>
<tr>
<td>Rhythmicm</td>
<td>–</td>
<td></td>
<td>–</td>
<td>57</td>
<td></td>
</tr>
<tr>
<td>Low in intelligencen</td>
<td>–</td>
<td></td>
<td>–</td>
<td>46</td>
<td></td>
</tr>
<tr>
<td>Pooro</td>
<td>–</td>
<td></td>
<td>–</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>Criminalp</td>
<td>–</td>
<td></td>
<td>–</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>Hostileq</td>
<td>–</td>
<td></td>
<td>–</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>Loudr</td>
<td>–</td>
<td></td>
<td>–</td>
<td>40</td>
<td></td>
</tr>
</tbody>
</table>

Note: The 10 most frequently selected traits in each study are in bold (9 or 11 traits are bolded in the case of incomplete information or ties). Unknown values, because of selective reporting in some of the studies, are indicated with a dash.

aAdditional trait reported by Gilbert (1951).
bTraits needed by Karlin et al. (1969) to account for the 10 most frequent selections.
cTraits added by Dovidio and Gaertner in 1982 (Dovidio & Gaertner, 1986).
dTraits needed to account for the 10 most frequent selections in the present study.
Gaertner, 1986), and the present study. In keeping with tradition (see Karmins et al., 1969), this table presents (a) the 12 traits most frequently selected in Katz and Braly's 1933 study, (b) the additional trait reported by Gilbert in 1951 (marked with a superscript "a"), (c) the traits needed by Karmins et al. in 1969 to account for the 10 most frequent selections (marked with a superscript "b"), (d) the traits added by Dovidio and Gaertner in 1982 (marked with a superscript "c"), and (d) the traits needed to account for the 10 most frequent selections in the present study (marked with a superscript "d"). Unknown values, because of selective reporting in some of the studies, are indicated with a dash. The 10 most frequently selected traits in each study are in bold (9 or 11 traits are bolded in the case of incomplete information or ties).

Perhaps the most striking feature of the data in Table 1 is the fact that 6 of the 10 most frequently selected adjectives in the present study were new additions to Katz and Braly's (1933) list. These 6 adjectives were added because they were found to be an integral part of the contemporary Black stereotype in a recent free-response assessment (Devine, 1989, Study 1). Katz and Braly's list was also based on then-recent free-response data; therefore, it is highly probable that their results accurately and somewhat comprehensively captured the then-contemporary stereotype of Blacks. In comparing the present data with those obtained by Katz and Braly, it is clear that the content of the Black stereotype has indeed undergone revision over the years, and it is possible to delineate the precise nature of this content change. Subsequent Princeton trilogy studies (and the 1982 Dovidio and Gaertner study, cited in Dovidio & Gaertner, 1986) were not based on then-recent free-response data; these investigations merely employed the same list of 84 adjectives used in the initial Katz and Braly assessment. Therefore, although the data from these studies clearly demonstrated a change in stereotype content, the precise nature of this change remains an open question. That is, it is likely that traits representing the then-contemporary stereotype were not present on Katz and

Braly's list, forcing respondents to select from an incomplete offering of adjectives.

Stereotype consistency is commonly represented by a stereotype uniformity index, which is indicated by the smallest number of traits needed to account for 50% of the total number of trait selections. The stereotype uniformity indexes for the Princeton trilogy studies and for the present investigation are reported in Table 4.2. Although the second and third Princeton trilogy studies appear to support the proposition that the stereotype was becoming less consistent over the years, the present uniformity index is nearly exactly that obtained in the original Princeton trilogy study. In light of this new data point, it is possible to offer an alternative to the fading stereotype proposition. It may be that despite the change in stereotype content over the years, there has remained a coherent, consistent stereotype of Blacks. Providing respondents an outdated, inadequate list of adjectives in the second and third Princeton trilogy studies may have spuriously elevated the uniformity indexes, thereby ostensibly representing fading (decreased consistency). Specifically, faced with an incomplete set of adjectives from which to choose, respondents may have resorted to selecting synonyms or simply less adequate traits, resulting in a less reliable assessment and an increase in the uniformity index.

Favorability ratings (in standard deviation units) were obtained from Rothbart and Park's (1986) comprehensive summary of several characteristics of a large list of trait adjectives. Based on these ratings, favorability indexes were computed for each sample that reflected the average degree of favorability of the most frequently selected traits. Table 4.2 displays the favorability indexes for the

3Despite the fact that the present sample, like Dovidio and Gaertner's (1986), was not drawn from Princeton University, we believe that students at the University of Wisconsin and Princeton University are similar enough in the attributes relevant to racial attitudes to warrant the comparisons offered in this section.

4Rothbart and Park (1986) did not report favorability ratings for several of the adjectives that were added to Katz and Braly's (1933) list in the present study (e.g., athletic, criminal, poor, rhythmic). To obtain favorability ratings for these traits, we repeated Rothbart and Park's procedure for assessing the favorability of traits for the new words and for a random subset of 20 of Katz and Braly's traits. A total of 25 respondents completed the favorability rating task. Our results for Katz and Braly's traits replicated Rothbart and Park's findings, in both the direction and the magnitude of the favorability scores. Moreover, the favorability ratings for the new words also appeared to be consistent with Rothbart and Park's findings in terms of both direction (e.g., athletic and rhythmic were positive, whereas poor and criminal were negative) and magnitude (e.g., criminal was rated as more negative than poor).
to select from an index, which is a commonly represented index, which is a traits needed number of trait indexes for this for the present le. 4.2. Although ilogy studies ap- at the stereotype per the years, the rly exactly that this study. In possible to offer type proposition. ge in stereotype remained a colacks. Providing a late list of adjec- cet trilogy study the uniformity essentially fading rally, faced with from which to sorted to selecte- rate traits, re- it and an increase

TABLE 4.2. Uniformity and Favorability Indexes for the Various Stereotype Assessment Studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Uniformity</th>
<th>Favorability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Katz and Braly (1933)</td>
<td>4.6</td>
<td>-0.60</td>
</tr>
<tr>
<td>Gilbert (1951)</td>
<td>12.0</td>
<td>-0.36</td>
</tr>
<tr>
<td>Karlin, Coffman, and Walters (1969)</td>
<td>12.3</td>
<td>+0.02</td>
</tr>
<tr>
<td>The 1982 Dovidio and Gaertner study (cited in Dovidio &amp; Gaertner, 1986)</td>
<td>NA</td>
<td>+0.21</td>
</tr>
<tr>
<td>The present study</td>
<td>4.5</td>
<td>-0.64</td>
</tr>
</tbody>
</table>

Note: NA = not available.

Princeton trilogy studies, the 1982 Dovidio and Gaertner investigation (cited in Dovidio & Gaertner, 1986), and the present study. These data appear to represent a peculiar pattern; it seems that the stereotype becomes progressively less negative over the years until the present study, at which time the favorability index reverts back to the negativ-ity of the 1933 sample.

We believe this pattern is only peculiar if one assumes that the previous studies were measuring (as they purport) the Black stereotype. Although the ambiguity of instructions given to respondents precludes any definitive statements, if the previous stereotype assessment studies actually represent (primarily) the measurement of respondents personal beliefs about Blacks, the data are no longer peculiar. The present study, in this scenario, is the only one of the assessments under consideration that explicitly measured the Black stereotype, and the data from our study suggest the existence of a consistent and highly negative contemporary stereotype of Blacks. We now turn to the personal beliefs data to determine if these data are congruent with an interpretation of the previous studies as (primarily) personal beliefs assessments.

Personal Beliefs Assessment

Although none of our respondents refused to complete the stereotype assessment task, 21% (all but one of whom were low prejudiced) declined the request to fill out the personal beliefs task. Interestingly, both Gilbert (1951) and Karlin et al. (1969) also reported that a number of their respondents refused to complete the task presented to them. Respondents in our study who refused to complete the personal beliefs assessment gave reasons that were very similar to those expressed by the respondents in the studies of Gilbert and Karlin et al. (e.g., I don’t think you can characterize a whole group of people with a list of traits—everyone is an individual). This differential pattern of refusals for the stereotype and personal beliefs assessments lends credence to our argument that the Princeton trilogy studies primarily assessed personal beliefs.

As with the stereotype assessment results, our initial consideration will be an examination of the frequency of traits chosen to represent respondents’ personal beliefs and the consistency with which these traits were selected. As with the stereotype assessment, a substantial number of the most frequently selected traits (4 of 10) were new additions to Katz and Braly’s (1933) list (e.g., athletic, rhythmic, poor). These data underscore the importance of providing respondents with an updated and comprehensive list when assessing their representations of social groups. Computation of a personal beliefs uniformity index further attests to the importance of a complete set of adjectives. The personal beliefs uniformity index was 4.6, suggesting that respondents’ personal beliefs about Blacks were consistent in their organization.

Of central importance in determining whether the previous assessments measured stereotypes or personal beliefs is the personal beliefs favorability index. Computation of this index from Rothbart and Park’s (1986) ratings yielded a value of +.22. Juxtaposing the present favorability index with those from previous studies (see Table 4.2) strongly suggests that the previous studies were primarily assessing personal beliefs, not stereotypes. Adopting this interpretation enables one to conclude that personal beliefs toward Blacks have progressively become more favorable over the years, to the point that they are, at present, predominately positive. Therefore, it appears that the Princeton trilogy is relatively silent on the issue of stereotype fading but speaks clearly, and somewhat optimistically, to the issue of personal belief change.

High- Versus Low-Prejudiced Individuals

In addition to making interstudy comparisons, we also sought to examine the effects of respondents’

---

3 The personal beliefs favorability index obtained in the present study is most likely a conservative value, given that ratings from the refusals (who presumably possessed predominantly positive personal beliefs) could not be incorporated into the computation of the index.
prejudice level on their knowledge of and belief in the Black stereotype. The total number of traits selected by respondents was submitted to a mixed-model analysis of variance (ANOVA), with prejudice level (high vs. low) as the between-subjects variable and type of assessment (stereotype vs. personal beliefs) as the within-subjects variable. This analysis yielded a type of assessment main effect, \( F(1, 112) = 26.69, p < .001 \), indicating that respondents selected more traits during the stereotype (\( M = 19.96 \)) than during the personal beliefs (\( M = 15.29 \)) assessment.

Columns 2 and 3 of Table 4.3 display the traits most frequently selected by low- and high-prejudiced respondents to characterize the Black stereotype; columns 4 and 5 display the comparable frequency data for respondents’ personal beliefs. The 10 most frequently selected traits within each column are in bold print (11 in the case of ties). For each trait in Table 4.3, the percentage of respondents selecting that trait was submitted to a Prejudice Level \( \times \) Type of Assessment mixed-model ANOVA. As indicated in the table, 15 of the 24 interactions (which test the effect of central interest) were significant (\( p < .05 \)), and another 4 attained a marginal level of significance (\( p < .10 \)).

Decomposition of the 19 significant and marginally significant interactions revealed that high-prejudiced respondents’ personal beliefs significantly (\( p < .05 \)) diverged from their knowledge of the stereotype for only 9 of the 19 traits, whereas the personal beliefs and knowledge of the stereotype for low-prejudiced respondents differed significantly (\( p < .05 \)) for all 19 of the traits considered (see Table 4.3). All of these effects were in the anticipated direction (i.e., the percentage for negatively valenced traits was greater in the stereotype than in the personal beliefs assessment and vice versa for positively valenced traits). A second set of comparisons probed for prejudice-level differences within assessment type. These analyses revealed no significant differences for the stereotype assessment, whereas all 19 prejudice-level comparisons for the personal beliefs assessment attained significance (\( p < .05 \)). Again, all these effects were in the anticipated direction (i.e., low-prejudiced respondents reported a greater percentage than high-prejudiced respondents for the positive traits and vice versa for the negative traits).

Together, these comparisons clearly support Devine’s (1989) hypothesis that high- and low-prejudiced individuals both possess the same stereotype of Blacks but that the stereotype is only endorsed by the former group of individuals.

Respondents’ average favorability scores were analyzed in a Prejudice Level \( \times \) Type of Assess-
ment mixed-model ANOVA. This analysis yielded a main effect for prejudice level, $F(1, 112) = 50.90$, $p < .001$, indicating that across assessments, high-prejudiced respondents’ trait selections were less favorable ($M = .11$) than those for low-prejudiced respondents ($M = .03$). A type of assessment main effect, $F(1, 112) = 228.16$, $p < .001$, was also obtained, such that respondents’ overall trait selections were less favorable for the stereotype ($M = .15$) than for the personal beliefs ($M = .01$) assessment. Importantly, both of these main effects were qualified by a significant Prejudice Level x Type of Assessment interaction, $F(1, 112) = 58.18$, $p < .001$. The means clearly show that high- and low-prejudiced respondents did not select differentially valenced traits for the stereotype (both $M_s = .15$), whereas the two prejudice levels diverged sharply on the personal beliefs assessment ($M_s = .07$ and $+.09$ for high- and low-prejudiced respondents, respectively, $p < .001$). Analyses employing the number of negative and positive traits selected in the two assessments both yielded results identical to those for the average favorability scores (i.e., the same two main effects and interactions were significant) and highly comparable patterns of means.

Discussion

It is often argued that stereotype change is an important precursor to prejudice reduction and the amelioration of intergroup relations. As such, it is important to carefully evaluate the techniques used to assess the nature and extent of such change. Classic stereotype assessment studies were found to be lacking in several ways that encourage caution with regard to statements concerning stereotype stability or change. First, historically, stereotype assessment research has been guided more by methodological factors than theoretical and conceptual considerations (Ashmore & Del Boca, 1981). In the present study, we applied Devine’s (1989) conceptual distinction between stereotypes and personal beliefs to the classic Princeton trilogy studies in the interest of reevaluating the oftespoused claim that the Black stereotype is changing in content, consistency, and valence (Brehm & Kassin, 1990; Dovidio & Gaertner, 1986). Second, we addressed several methodological shortcomings of the classic adjective checklist procedure that prevented clear conclusions about stereotype change. By disambiguating the instructions provided to respondents, attaining an independent assessment of prejudice, and presenting respondents with an up-to-date and comprehensive list of trait adjectives, we were able to obtain a precise snapshot of our respondents’ representations of Black Americans as a social group. This picture leads us to the conclusion that the Black stereotype is not fading among Whites; rather, personal beliefs about Blacks are undergoing a revision.

Results from our stereotype assessment suggest that there is a clear, consistent contemporary stereotype of Blacks and that this stereotype is highly negative in nature. Our personal beliefs data, however, indicate that there is an equally consistent set of beliefs about Blacks that, in stark contrast to the contemporary stereotype, is positively valenced. The Princeton trilogy assessments did not distinguish between stereotypes and personal beliefs but merely asked respondents to select traits that seemed typical of Blacks. From the comments provided Gilbert (1951) and Karins et al. (1969), it is clear that some construed the task as a stereotype assessment, whereas others thought their personal beliefs were being assessed. This variability in task construal makes it difficult to make definitive statements about the Princeton trilogy data.

In spite of this interpretational obstacle, we believe that a comparison of the data pattern from the Princeton trilogy studies (and the 1982 Dovidio and Gaertner study, cited in Dovidio & Gaertner, 1986) to that obtained in the present investigation strongly suggests that the previous studies primarily assessed respondents’ personal beliefs. The favorability ratings across the previous studies display a clear trend from highly negative scores in the 1933 sample ($-.60$) to a more positive value in Dovidio and Gaertner’s sample ($+.21$). Interestingly, the favorability index for personal beliefs in the present study ($+.22$) fits nicely into the temporal trend observed in the previous investigations, whereas the favorability ratings for the stereotype ($-.64$) matches that of the initial Katz and Braly (1933) assessment. We believe these data indicate that respondents primarily reported their personal beliefs in the traditional Katz and Braly task, beliefs that were predominantly negative (and, presumably, highly congruent with the then-contemporary stereotype) in 1933 and progressively became more positive as a result of changes in the
fabric of American society (see Dovidio & Gaertner, 1986; Schuman, Steeth, & Bobo, 1985, for reviews). This interpretation is bolstered by the fact that none of Katz and Braly’s respondents refused to complete the task (as with our stereotype assessment), whereas a number of respondents refused to complete the task in the later Princeton trilogy studies (as with our personal beliefs assessment). Thus reevaluating the Princeton trilogy studies in light of our present investigation impels us to conclude that the classic stereotype assessment studies are not stereotype assessment studies at all; rather, these studies assessed personal beliefs, which may (in Katz & Braly, 1933) or may not (in all the subsequent studies) be congruent with the stereotype.

Our inclusion of an independent measure of prejudice, like our independent assessments of the stereotype and personal beliefs, enabled us to obtain a more precise understanding of respondents’ representations of Black Americans. Both high- and low-prejudiced respondents apparently possess the same stereotypic knowledge in their associative network, as the two groups did not differ in their reports of the content or valence of the Black stereotype. In contrast, high- and low-prejudiced respondents diverged sharply in reports of their personal beliefs. High-prejudiced respondents endorsed beliefs that overlapped considerably with their knowledge of the Black stereotype, whereas the beliefs of low-prejudiced respondents were highly discrepant from their knowledge of the stereotype. These data provide strong support for Devine’s (1989) proposition that high- and low-prejudiced individuals alike possess equivalent knowledge of the Black stereotype; it is in their endorsement of the stereotype that they diverge. Thus the present study not only replicates Devine’s stereotype assessment findings but also provides the first direct evidence in support of the personal beliefs aspect of Devine’s model. More important, assessing both the stereotype and personal beliefs in the same study enabled us to empirically validate the stereotype/personal-beliefs distinction in compelling fashion—by treating type of assessment as a within-subjects variable.

It is interesting to note that the low-prejudiced respondents in our sample may be divided into two types: those who agreed to complete the stereotype but not the personal beliefs assessment and those who agreed to complete both tasks. Those in the first group apparently perceived the very process of ascribing traits to Blacks when reporting personal beliefs as an unacceptable activity and refused to engage in the task. Those in the second group agreed to complete the task but resorted to selecting positive, often counterstereotypic traits (e.g., intelligent, honest) to represent their personal beliefs. Thus, although they differed in strategy, both types of low-prejudiced respondents were unwilling to endorse negative stereotypic generalizations about Blacks.7 High-prejudiced respondents, in contrast, reported personal beliefs that overlapped considerably with their knowledge of the stereotype.

Consideration of the stereotype/personal-beliefs distinction is essential in attempting to understand the social psychological experience of the low-prejudiced individual. The coexistence of a rejected yet enduring negative stereotype of Blacks and a positive set of beliefs about Blacks places the low-prejudiced social perceiver in a precarious position. A considerable amount of research indicates that stereotypes can be automatically activated by the perception of social stimuli (e.g., exemplars of the group, group labels), resulting in prejudice-like feelings, thoughts, and behaviors for high- and low-prejudiced individuals alike (Devine, 1989; Klinger & Beall, 1992). It is only when low-prejudiced individuals possess the time and cognitive capacity necessary to engage in controlled processing that they are able to inhibit the automatically activated stereotype and respond based on their personal beliefs (Devine, 1989; Klinger & Beall, 1992). Functional eradication of the stereotype from the low-prejudiced individual’s associative network is possible (Monteith, 1993) but is likely to represent an arduous and prolonged task (Devine & Monteith, 1993). In the process of “breaking the prejudice habit,” the low-prejudiced individual is susceptible to negative self-directed

Although the refusal of a large portion of our low-prejudice respondents to complete the personal beliefs task would typically make comparisons involving prejudice level and/or the personal beliefs assessment tenuous, in the present study this attrition actually made for a more conservative set of analyses. Low-prejudiced respondents who agreed to complete the personal beliefs assessment primarily selected positive traits to characterize their personal beliefs about Blacks. It is almost certain that those who refused to complete the personal beliefs assessment would have adopted a comparable, and perhaps even more extreme, version of this strategy. Therefore, the absence of these respondents actually has the effect of minimizing the probability that we would find prejudice-level differences in our analyses.
affect (Devine, Monteith, Zuwerink, & Elliot, 1991) when their feelings, thoughts, or behaviors prove discrepant from their internalized personal beliefs or standards (Monteith, Devine, & Zuwerink, 1993).

Our critique of the stereotype assessment literature in the present study focused on Katz and Braly’s (1933) technique because it has been the workhorse of stereotype measurement over the years. Many of the alternative stereotype assessment devices that have been developed possess the same weaknesses that we have identified in Katz and Braly’s procedure, the most important being the failure to clearly distinguish between stereotypes and personal beliefs (e.g., Brigham, 1971; Gardner, Lalonde, Nero, & Young, 1988; McCauley & Stitt, 1978). Only those measures that assess stereotypic associations nonreactively (e.g., Dovidio et al., 1986; Gaertner & McLaughlin, 1983) appear to be beyond the scope of our critique. Despite their many positive qualities, however, nonreactive measures are somewhat limited in their use, as they cannot be employed to assess controlled processes such as personal beliefs (see Klinger & Beall, 1992).

Besides our own interpretation of the Princeton trilogy data, there is an additional explanation that deserves comment. Some have suggested that changes in the pattern of traits selected in assessment studies over the years simply reflect respondents’ increasing sensitivity to nonprejudiced social norms rather than true stereotype or belief change (Crosby, Bromley, & Saxe, 1980; Dovidio & Gaertner, 1991; Sigall & Page, 1971). Clearly, this social desirability argument is not applicable to our stereotype assessment data, because these data revealed the existence of a contemporary stereotype that is highly consistent and negative in valence. High-prejudiced respondents’ reports of their personal beliefs overlapped considerably with their knowledge of the stereotype; therefore, their pattern of trait selection cannot be attributable to social desirability concerns either. Thus a social desirability interpretation of the present data requires the assumption that low-prejudiced respondents—the only group that reported predominantly positive personal beliefs about Blacks—are more sensitive to social desirability demands than their high-prejudiced counterparts.

Although we cannot definitively rule out the possibility that our results for low-prejudiced respondents reflect social desirability demands, we think it is highly unlikely for the following three reasons. First, our procedure was carefully designed to minimize social desirability concerns; all respondents completed both the stereotype and the personal beliefs assessments under conditions that highlighted anonymity. Second, research by Monteith (1993) has demonstrated that behavior discrepant from the personal standards of low-prejudiced respondents creates negative self-directed affect and activates prejudice reduction processes in these individuals. Third, in related research, we have found that for those who report low-prejudiced attitudes toward Blacks, responding to Blacks in a nonprejudiced manner is a highly important, internalized aspect of their self-conceptions (Devine et al., 1991; Zuwerink et al., 1996). This research has further suggested that it is high-prejudiced, rather than low-prejudiced, respondents who are highly sensitive to social desirability cues (see Devine et al., 1991, Study 3).

Results from the present study validate our methodological critique of the traditional Katz and Braly (1933) procedure and provide a clearer picture of the complex nature of Whites’ representations of Blacks. We recommend that researchers interested in employing this assessment procedure in future research (a) provide respondents with unambiguous instructions that indicate whether they should select traits based on their knowledge of the stereotype or based on their personal beliefs about the target group, (b) include an independent assessment of prejudice if respondents’ personal beliefs about the target group are of interest, and (c) provide respondents with a list of adjectives that has the potential to comprehensively capture all aspects of the contemporary stereotype and/or set of beliefs about the target group (this requires that the adjective list be based on recent free-response data). Implementing these modifications in the present study yielded a data set that provided us with a clear picture of respondents’ cognitive representations of Blacks. Therefore, in spite of our critique of Katz and Braly’s procedure, we believe that it can, in modified form, be an effective tool in contemporary stereotype and prejudice research.

Our data suggest an apparent paradox between considerable belief change (at least among some respondents) and stereotype persistence. If stereotypes are ceasing to represent what people believe about a social group, why do they persist? A number of individual-level cognitive and sociocultural
factors contribute to the persistence of stereotypes. First, it is important to recognize that individual-level changes in one’s beliefs (i.e., renouncing prejudice and stereotypes) does not lead to the immediate elimination of stereotypes from memory (Devine, 1989; Higgins & King, 1981). Moreover, during the process of prejudice reduction, stereotype-based responses are highly accessible and serve as rivals to belief-based responses (Devine, 1989; Jamieson & Zanna, 1989; Kruglanski & Freund, 1983). Thus change at the individual level is not all-or-none; overcoming stereotyped-based responding requires a great deal of personal motivation and effort (Devine, 1989; Monteith, 1993). An additional challenge to such efforts is suggested by recent evidence that efforts to suppress stereotypes may even heighten their accessibility (Macrae, Bodenhausen, Milne, & Jetten, 1994).

One of the main impediments to the fading of racial stereotypes is that they remain deeply embedded in the cultural fabric of our nation (Ehrlich, 1973; Jones, 1972). Despite a shift in social norms regarding overt expressions of prejudice and discrimination, stereotypic images of Blacks persist in the dominant media (e.g., television, newspapers), and Blacks continue to be underrepresented in traditional positions of power (e.g., education and industry; Schuman et al., 1985). As a result, stereotypes are perpetuated within the culture in subtle, yet highly effective, ways. Although efforts such as affirmative action and increased sensitivity to media portrayals of Blacks may ultimately have positive effects and facilitate the fading of cultural racial stereotypes, it is likely that such changes will take place over a protracted period of time. We, optimistically, believe that change will eventually take place; our analysis of stereotypes simply leads us to the conclusion that this change, at both the macro (societal) and micro (individual) levels, takes time to unfold.

REFERENCES


