The Rules of Implicit Evaluation by Race, Religion, and Age

Jordan R. Axt, Charles R. Ebersole and Brian A. Nosek

Psychological Science 2014 25: 1804 originally published online 30 July 2014
DOI: 10.1177/0956797614543801

Research Article

Psychological Science
2014, Vol. 25(9) 1804–1815

DOI: 10.1177/0956797614543801
pss.sagepub.com

The Rules of Implicit Evaluation by Race, Religion, and Age

Jordan R. Axt[1], Charles R. Ebersole[1], and Brian A. Nosek[1,2]
[1]Department of Psychology, University of Virginia, and [2]Center for Open Science, Charlottesville, Virginia

Abstract

The social world is stratified. Social hierarchies are known but often disavowed as anachronisms or unjust. Nonetheless, hierarchies may persist in social memory. In three studies (total $N > 200{,}000$), we found evidence of social hierarchies in implicit evaluation by race, religion, and age. Participants implicitly evaluated their own racial group most positively and the remaining racial groups in accordance with the following hierarchy: Whites > Asians > Blacks > Hispanics. Similarly, participants implicitly evaluated their own religion most positively and the remaining religions in accordance with the following hierarchy: Christianity > Judaism > Hinduism or Buddhism > Islam. In a final study, participants of all ages implicitly evaluated age groups following this rule: children > young adults > middle-age adults > older adults. These results suggest that the rules of social evaluation are pervasively embedded in culture and mind.

Keywords

social cognition, racial and ethnic attitudes and relations, social perception, prejudice, associative processes, open data, open materials

Received 1/2/14; Revision accepted 6/22/14

Social status is relational: Some people are higher status, and others are lower status. Although such social-status hierarchies are often deemed antiquated or unfair, this differential status both reflects and causes differential outcomes for groups. Groups with higher status enjoy superior academic outcomes (Sirin, 2005), perceive less discrimination against their own group (Kessler, Mickelson, & Williams, 1999), report better physical and mental health (Williams, Yu, Jackson, & Anderson, 1997), create broader social networks (Campbell, Marsden, & Hurlbert, 1986), and receive more opportunities and consideration for coveted positions (Lin, Ensel, & Vaughn, 1981).

If hierarchies are consensual, then they may become cultural truisms that sustain differential assessments of who is valued and create differential opportunities and outcomes across social groups (Sidanius & Pratto, 1999). If status hierarchies are idiosyncratic—for example, if each group perceives itself to be on top—then in-group favoritism may sustain intergroup conflict in the competition for opportunities and resources (Hagendoorn & Hraba, 1987; Hraba, Hagendoorn, & Hagendoorn, 1989; Tajfel, 1982).

There is support for both of these possibilities. In a variety of cultures, both modern and historical, judgments of which groups are higher and lower status appear to reflect consensus, particularly when the judgments focus on social power, one element of status. For example, consensus has been shown among ethnic groups in the former Soviet Union (Hagendoorn, Drogendijk, Tumanov, & Hraba, 1998), Canadian immigrants (Berry & Kalin, 1979), ethnic youth in the Netherlands (Verkuyten, Hagendoorn, & Masson, 1996), and ethnic groups in modern American society (Kahn, Ho, Sidanius, & Pratto, 2009).

Status is partly a function of social power, but is also a function of social evaluation. Some groups are evaluated more favorably than others, and these judgments are distinct from judgments of social power. On the one hand, there is substantial evidence for in-group favoritism

Corresponding Author:
Jordan R. Axt, Department of Psychology, University of Virginia, Box 400400, Charlottesville, VA 22904-4400
E-mail: jaxt@virginia.edu

among both high- and low-status groups (Bettencourt, Dorr, Charlton, & Hume, 2001; Mullen, Brown, & Smith, 1992). According to social identity theory, humans have a desire to maintain a positive social identity, which they often achieve by favoring their own group and by believing that group to be superior to others (Tajfel, 1978). These in-group biases are particularly pronounced when group boundaries are believed to be stable and relatively impermeable (Tajfel & Turner, 1986).

On the other hand, system-justification theory suggests that there are limits to in-group favoritism when it is clearly at odds with the existing status hierarchy—the "status status quo" (Jost, Banaji, & Nosek, 2004). According to system-justification theory, there is a pervasive tendency to see the world as just and fair, even if one's own group is not atop the hierarchy. This tendency is particularly evident in measures of implicit social cognition (Jost et al., 2004). Even when people do not consciously endorse hierarchies, they may nonetheless learn and encode them in social memory. Such associations may be the basis for automatic responses that shape perception, judgment, and action (Nosek, Hawkins, & Frazier, 2011, 2012). In sum, status hierarchies may be evident in implicit social evaluation, even among groups with low status and despite tendencies for in-group favoritism.

Using American samples, we investigated the presence of hierarchies in social evaluation—who is good—in three domains: race, religion, and age. We further examined whether these hierarchies were consensually shared among social groups, and whether they showed evidence of in-group favoritism. Finally, we examined hierarchies in both explicit and implicit social cognition, hypothesizing that they would be particularly likely to be observed in implicit social cognition (as markers of cultural influence) despite whatever people may believe and endorse consciously. Our results provide support for the existence of pervasive hierarchies in social evaluation that complement evidence for pervasive hierarchies in social power.

Study 1

Method

Participants. The 97,641 participants in this study completed at least one measure while the study was the featured task at Project Implicit (implicit.harvard.edu), from June 5, 2012, to April 11, 2013.[1] The end date of the study was determined arbitrarily when another task became the featured task. Given the possible cultural specificity of the rules we were examining, only American citizens or residents were included in the analyses (81.3% of participants who reported demographic information). The ordinal pattern of results did not change when we included all participants; results from the full sample are included in the Supplemental Material available online (Tables S4–S7).

Among participants reporting demographic information, 61% were female, and the mean age was 30.2 years (SD = 13.5). By race, 70.8% were White, 8.5% Black, 3.2% East Asian, 2.0% South Asian, and 5% biracial; 10.5% indicated other or unknown racial membership. By ethnicity, 10.1% of participants who provided demographic information (n = 5,351) were Hispanic or Latino and were included in the Hispanic group for analyses; non-Hispanic participants were classified as White (n = 37,314), Black (n = 4,514), East Asian (n = 1,756), or of another or unknown (according to self-report) race (n = 4,216).

Procedure. The study session consisted of four components, completed in a random order: two surveys that were not analyzed for this report, a demographics questionnaire followed by a survey on racial attitudes, and a four-category race Multicategory Implicit Association Test (MC-IAT). After completing all measures, participants were debriefed and given feedback on their MC-IAT performance (see https://osf.io/zg2su/ for demonstration links to view the study protocols for all studies).

Demographic and racial surveys. The demographics questionnaire included 15 items, but we analyzed only the data relating to race, ethnicity, gender, and age. The survey of racial attitudes included 6 items concerning preferences for Black, White, Asian, and Hispanic people; for each of the possible pairs of these groups, participants indicated the degree to which they preferred one group over the other (7-point scale ranging from 1, *I strongly prefer* X *people to* Y *people*, to 7, *I strongly prefer* Y *people to* X *people*).

MC-IAT. The MC-IAT, a variant of the Brief IAT (Sriram & Greenwald, 2009), measured the strength of associations between racial groups and positive evaluation. The test contained 14 blocks, of which the first 2 were practice. In each block, items were presented one at a time, and participants categorized them as quickly as possible. Categorization errors had to be corrected before continuing to the next trial. In the first block (16 trials), participants pressed the "I" key for all "good" words (*Love, Pleasant, Great*, and *Wonderful*) and the "E" key for "other words" (*Hate, Unpleasant, Awful*, and *Terrible*). In the second block (20 trials), participants pressed the "I" key for all good words and for faces belonging to a specified racial group (Asian, Black, Hispanic, or White; stimuli were two male and two female faces, each shown above a prototypical surname, e.g., "N. Chang"[2]),

and they pressed the "E" key for "any other images and words" (the same negative words as in the first block and faces from one of the other three racial groups). For the remaining 12 blocks (16 trials each), the structure was the same as in the second block, with each block using a different combination of target and other racial group. For example, there were 3 blocks for which participants pressed the "I" key for Asian faces: In 1 block, the other faces were Black; in another, they were Hispanic; and in another, they were White. Randomization of block order was constrained so that each racial group appeared as a target once every 4 blocks. Participants were randomly assigned to 1 out of 24 possible block orders.

MC-IAT *D* scores were calculated following the guidelines outlined for the Brief IAT in Nosek, Bar-Anan, Sriram, Axt, and Greenwald (2014). This MC-IAT produced six *D* scores, representing each contrast of racial groups (White vs. Black, White vs. Asian, White vs. Hispanic, Asian vs. Black, Asian vs. Hispanic, Black vs. Hispanic).

To calculate each *D* score, we removed all trials with a response time greater than 10,000 ms, as well as the first four trials of each block, as these were practice. Next, all response times lower than 400 ms were recoded to 400 ms, and all response times greater than 2,000 ms were recoded to 2,000 ms. Participants' MC-IAT data were excluded if more than 10% of their response times were less than 400 ms, an indication of careless responding (2.9% of participants who completed the MC-IAT).

A *D* score was computed for each contrast by subtracting the mean latency for one block (e.g., White faces categorized with good words, Black faces categorized with bad words) from the mean latency for the other block (e.g., Black faces categorized with good words, White faces categorized with bad words) and then dividing by the standard deviation of the latencies across both blocks.

From these six contrast *D* scores, we computed an aggregate score for each race (e.g., the White score was the average of three *D* scores, i.e., scores comparing White with Asian, Black, and Hispanic people). This analysis strategy provided an evaluation of each group in comparison with the others. The four aggregate scores are interdependent; knowing three scores directly implies the fourth, and the mean of the four scores is necessarily 0. Thus, positive scores indicate evaluations more favorable than the average evaluation across the four groups, and negative scores indicate evaluations more unfavorable than the average evaluation across the four groups.

Results

Sample sizes vary among the tests reported because of missing data. For participants of all racial and ethnic groups, the order of implicit racial preferences was the same (Fig. 1). Whites, Asians, Blacks, and Hispanics exhibited the most positive associations for their own racial group. In addition, their implicit evaluations of the remaining racial groups always placed White people first, followed by Asian, Black, and then Hispanic people, all pairwise ts(1467–31,567) > 4.06, all ps < .001, all ds > 0.06, average d = 0.2, except for the comparison between the Black aggregate and Hispanic aggregate variables among Asian participants, $t(1468) = 0.86$, $p = .394$. Participants who identified with racial groups other than the four target groups (e.g., American Indians, Pacific Islanders) showed the same ordinal pattern, all ts(3291–3303) > 4.2, all ps < .001, all ds > 0.08, average d = 0.10. Table 1 presents descriptive statistics.

In general, the individual group contrasts showed transitive relations with the aggregate scores. Thus, knowing only the aggregate relations among the four groups provides sufficient information to derive relatively accurate estimates for any particular pairing, and this held for the subsequent two studies as well. For example, Whites' preference between any two groups calculated by taking the difference between the two groups' aggregate scores was always within one fifth of a standard deviation of the corresponding contrast score. Thus, the six contrast scores do not provide much information in addition to what can be derived from the summary scores.

The ordinal relations among racial groups were less consistent for explicit evaluations than for implicit evaluations. Whites, Asians, Blacks, and Hispanics did show an explicit preference for their own racial group compared with the others, and most groups' explicit evaluations showed a hierarchy from Whites to Asians, Hispanics, and then Blacks. Note that for White and Asian participants, the explicit preference for Hispanic people over Black people reversed the order reflected in implicit evaluation. Another reversal was that Black participants explicitly preferred Hispanics to Asians and Whites, on average. The Supplemental Material provides *t* and *d* values for implicit evaluations (Table S1), descriptive statistics and *t* and *d* values for explicit evaluations (Tables S2 and S3), descriptive statistics for each of the six implicit and explicit contrasts (Tables S8 and S9), and correlations between implicit and explicit evaluations (Table S11).

Study 2

In Study 2, we tested whether invariance in implicit evaluation would be observed for another social identity: religion.

Method

Participants. The 353,048 participants in Study 2 completed at least one measure while this study was the

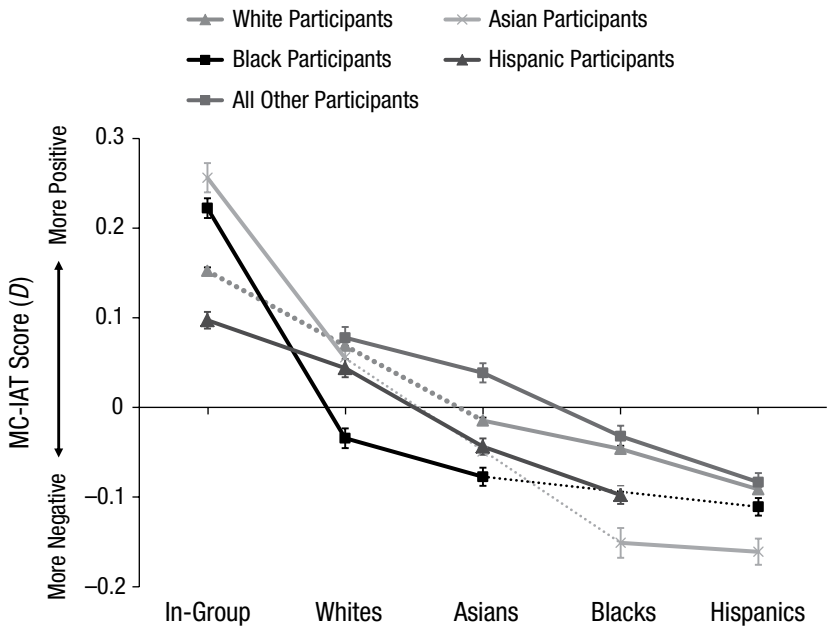

Fig. 1. Results from Study 1: Multicategory Implicit Association Test (MC-IAT) *D* scores for the in-group and White, Asian, Black, and Hispanic people. Results are shown separately for participants belonging to these four racial groups and for participants who indicated that they were of other races. Dotted lines connect means that skip over the in-group. Error bars denote 95% confidence intervals on the means.

featured task at Project Implicit, from June 20, 2009, to June 13, 2013.[3] The end date of the study was determined arbitrarily once data from at least 2,500 participants from each targeted religion group from the full sample had been collected. Only American citizens or residents (85.2% of participants reporting demographic information) were included in analyses. The ordinal pattern of results did not change when we included all participants; results from the full sample are included in the Supplemental Material (Tables S4–S7).

Among participants reporting demographic information, 58.5% were female, and the mean age was 28.3 years (*SD* = 12.8). By race, 76.4% were White, 6.9% Black, 2.4% East Asian, 2.5% South Asian, and 6.2% biracial; 5.7% reported that they were of other or unknown races. By ethnicity, 9.2% were Hispanic or Latino. For analyses, on the basis of self-reported religious affiliation, participants were classified as Christian (n = 98,909), Jewish (n = 8,754), Buddhist (n = 2,878), Hindu (n = 1,861), or Muslim (n = 3,295) or as belonging to another or no religion (n = 77,698).

Procedure. Each study session had three components, presented in a randomized order. Participants completed an attitudinal survey, a demographics questionnaire, and a four-category religion MC-IAT that compared evaluations

Table 1. Implicit Race Attitudes: Descriptive Statistics From Study 1

Participant's race	Group evaluated											
	White people			Asian people			Black people			Hispanic people		
	n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
White	31,656	0.15	0.36	31,684	–0.01	0.30	31,674	–0.05	0.33	31,664	–0.09	0.29
Asian	1,469	0.06	0.34	1,474	0.26	0.32	1,472	–0.15	0.32	1,472	–0.16	0.29
Black	3,676	–0.03	0.34	3,680	–0.08	0.31	3,677	0.22	0.34	3,676	–0.11	0.30
Hispanic	4,411	0.04	0.34	4,414	–0.04	0.31	4,413	–0.10	0.34	4,411	0.10	0.32
Other	3,307	0.08	0.35	3,311	0.04	0.32	3,317	–0.03	0.35	3,307	–0.08	0.30

Note: The table presents mean aggregate *D* scores (Greenwald, Nosek, & Banaji, 2003) from the Multicategory Implicit Association Test; more positive values indicate greater preference for members of the target group.

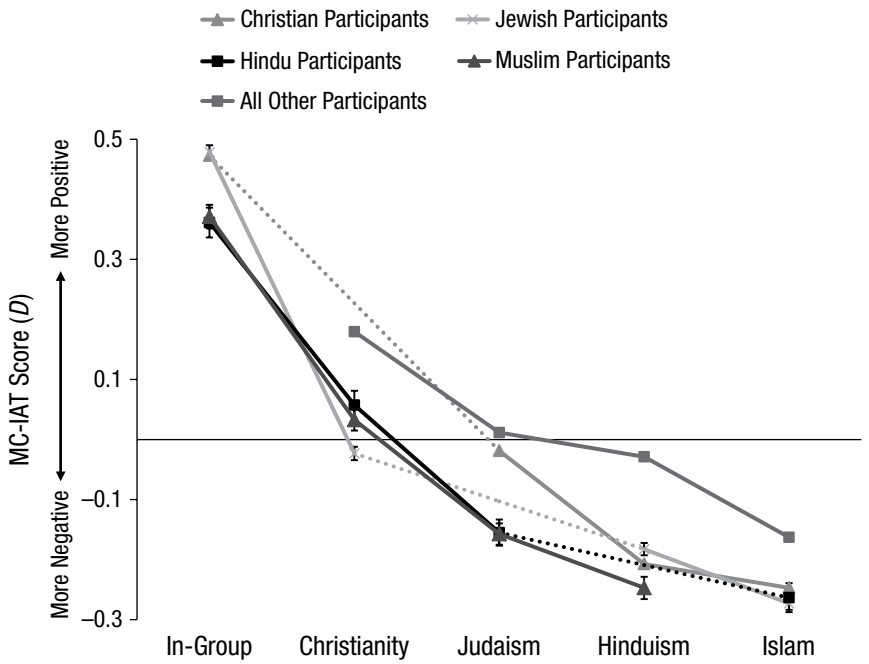

Fig. 2. Results from the Hinduism condition of Study 2: Multicategory Implicit Association Test (MC-IAT) *D* scores for the in-group and Christianity, Judaism, Hinduism, and Islam. Results are shown separately for participants belonging to these four religion groups and for participants who indicated that they had other religions. Dotted lines connect means that skip over the in-group. Error bars denote 95% confidence intervals on the means.

of Christianity, Judaism, Islam, and either Hinduism or Buddhism (subjects were randomly assigned to the two conditions).

Demographic questionnaire and attitudinal survey. The demographics questionnaire included 13 items, but we analyzed only the data relating to race, ethnicity, religion, gender, and age. Also, we analyzed responses to only 4 of the 16 items on the attitudinal survey. These items assessed perceptions of warmth toward each of the four religion groups included in the MC-IAT: "How warm or cold are your feelings toward *X*?" (rating scale from 1, *extremely cold*, to 9, *extremely warm*). The unanalyzed 12 items were random selections from a pool of 186 items about attitudes, beliefs, and ideology (Graham, Hawkins, & Nosek, 2012).

MC-IAT. The MC-IAT procedure was the same as in Study 1, but with religion categories and items used in place of categories and items representing race. The religion stimuli consisted of words associated with each religion: Christianity (*Gospel, Christian, Jesus, Church*), Islam (*Koran, Muslim, Muhammad, Allah*), Judaism (*Torah, Jew, Abraham, Yahweh*), Buddhism (*Karma, Buddhist, Buddha, Dharma*), and Hinduism (*Hindu, Krishna, Karma, Dharma*). Calculation of *D* scores, exclusion criteria, and the analysis strategy were the same as in Study 1 (data from 4.3% of participants who completed the MC-IAT were excluded).

Results

Sample sizes vary among the tests reported because of missing data. For participants of all religion groups, the order of implicit evaluations was the same (Figs. 2 and 3). Christians, Jews, Hindus, Buddhists, and Muslims exhibited the most positive associations for their own religion. In addition, their implicit evaluations of the remaining religions always placed Christianity first, followed by Judaism, Hinduism or Buddhism (depending on condition), and then Islam, all ts(729–40,517) > 3.85, all ps < .001, all ds > 0.09, average d = 0.44. Participants who did not belong to any of the target religions showed the same ordinal relations, all ts(32,882–33,085) > 11.1, all ps < .001, all ds > 0.06, average d = 0.22, with one exception: In the Buddhism condition, Buddhism was favored slightly over Judaism. Tables 2 and 3 present descriptive statistics.

Explicit evaluations did not show the same consistent ordinal relations. All five target religion groups reported an explicit preference for their own religion compared with the others. However, Jewish participants viewed Buddhism and Hinduism more warmly than Christianity. Muslim participants evaluated Buddhism more warmly than Judaism. Buddhist and Hindu participants evaluated

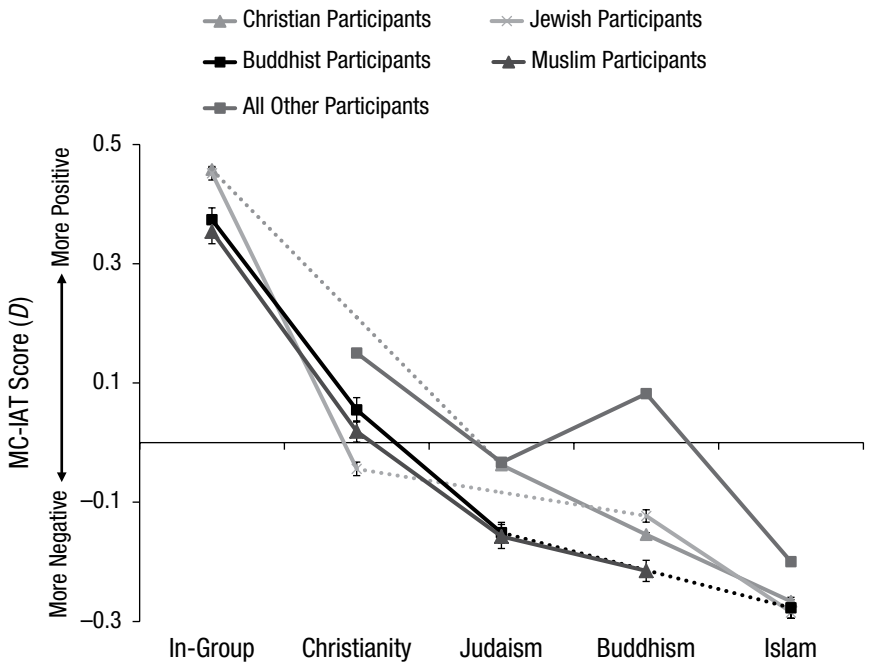

Fig. 3. Results from the Buddhism condition of Study 2: Multicategory Implicit Association Test (MC-IAT) *D* scores for the in-group and Christianity, Judaism, Buddhism, and Islam. Results are shown separately for participants belonging to these four religion groups and for participants who indicated that they had other religions. Dotted lines connect means that skip over the in-group. Error bars denote 95% confidence intervals on the means.

Judaism more warmly than Christianity. The Supplemental Material provides *t* and *d* values for implicit evaluations (Table S1), descriptive statistics and *t* and *d* values for explicit evaluations (Tables S2 and S3), descriptive statistics for each of the six implicit contrasts (Table S8), and correlations between implicit and explicit evaluations (Table S11).

Study 3

In Study 3, we tested whether invariance in implicit evaluation would be observed for one more social identity: age.

Method

Participants. The 49,014 participants in Study 3 completed at least one measure while this study was the featured task at Project Implicit, from April 21, 2011, to January 19, 2012.[4] The end date of the study was determined arbitrarily when another task became the featured task. The demographics questionnaire asked about citizenship but not residence, so only American citizens were included in analyses (76.6% of participants reporting demographics). The ordinal pattern of results did not change when we included all participants; results from

Table 2. Implicit Religion Attitudes: Descriptive Statistics From the Hinduism Condition of Study 2

Participant's religion	Religion evaluated											
	Christianity			Judaism			Hinduism			Islam		
	n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
Christianity	39,931	0.47	0.34	39,949	–0.02	0.30	39,935	–0.21	0.30	39,937	–0.25	0.29
Judaism	3,642	–0.03	0.33	3,644	0.48	0.34	3,639	–0.18	0.30	3,642	–0.27	0.31
Hinduism	732	0.06	0.32	735	–0.16	0.29	730	0.36	0.33	733	–0.26	0.32
Islam	1,286	0.03	0.31	1,281	–0.16	0.31	1,279	–0.25	0.32	1,277	0.37	0.34
Other or none	33,029	0.18	0.36	33,013	0.01	0.30	32,995	–0.02	0.32	33,007	–0.16	0.30

Note: The table presents mean aggregate *D* scores (Greenwald, Nosek, & Banaji, 2003) from the Multicategory Implicit Association Test; more positive values indicate greater preference for members of the target group.

Table 3. Implicit Religion Attitudes: Descriptive Statistics From the Buddhism Condition of Study 2

Participant's religion	Religion evaluated											
	Christianity			Judaism			Buddhism			Islam		
	n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
Christianity	40,661	0.46	0.34	40,659	–0.04	0.30	40,663	–0.15	0.31	40,666	–0.27	0.29
Judaism	3,642	–0.05	0.33	3,640	0.45	0.33	3,634	–0.12	0.31	3,639	–0.28	0.30
Buddhism	1,210	0.05	0.34	1,209	–0.15	0.29	1,209	0.37	0.33	1,205	–0.28	0.29
Islam	1,314	0.02	0.31	1,308	–0.16	0.30	1,306	–0.21	0.31	1,312	0.35	0.35
Other or none	33,212	0.15	0.36	33,222	–0.03	0.30	33,192	0.09	0.34	33,232	–0.20	0.31

Note: The table presents mean aggregate *D* scores (Greenwald, Nosek, & Banaji, 2003) from the Multicategory Implicit Association Test; more positive values indicate greater preference for members of the target group.

the full sample are included in the Supplemental Material (Tables S4–S7).

Among participants reporting demographic information, 68.7% were female, and the mean age was 33.3 years (*SD* = 14.1). By race, 72.5% were White, 10.3% Black, 2.3% East Asian, 1.5% South Asian, and 7.2% biracial; 6.2% reported that they were of other or unknown races. By ethnicity, 8.7% were Hispanic or Latino.

In the case of age, in-group status is not clearly delineated because there are few definitive markers for when one leaves one age group and enters another. This is nonconsequential for examining age attitudes because prior evidence suggests that age preferences are relatively steady across the age span, regardless of group membership (Nosek, Banaji, & Greenwald, 2002; Nosek et al., 2007). We report findings both using participant's age as a continuous variable and for age brackets: teens (*n* = 3,561), 20s (*n* = 10,113), 30s (*n* = 4,713), 40s (*n* = 3,965), 50s (*n* = 3,091), and 60s (*n* = 1,208).

Procedure. Each study session had three components, presented in a randomized order. Participants completed an attitudinal survey, a demographics questionnaire, and a four-category age MC-IAT.

Demographic and attitude surveys. The demographics questionnaire included nine items, but we analyzed only the data relating to race, ethnicity, gender, and age. The survey of age-group attitudes included four items, assessing attitudes toward "children," "young adults," "middle-aged adults," and "old adults," in that order. Participants were instructed to rate the warmth of their feelings toward each age group relative to the others (the age groups were represented by faces). They responded on an 11-point scale from 0 (*very cold*) to 10 (*very warm*; 5 = *neutral*). Participants also rated each age group's competence and likeability (on 8-point scales ranging from *extremely incompetent/unlikeable* to *extremely competent/likeable*).

MC-IAT. The four-category age MC-IAT had the same structure as the MC-IAT in Study 1 with minor changes. Test blocks had 18 instead of 16 critical trials, and there were six, rather than four, images for each age group. The word stimuli were the same as in Study 1. The age stimuli were selected on the basis of pretest ratings (*N* = 13) of the apparent age of faces taken from face databases (Ebner, 2008; Langner et al., 2010; Minear & Park, 2004) and of manufactured faces (generated using Fantamorph Deluxe software, www.fantamorph.com). Three male and three female stimuli, all White faces with neutral expressions and nondescript backgrounds, were selected for each of the four age groups (mean rated ages from the pretest are given in parentheses): children (*M* = 11.1 years, *SD* = 1.9), young adults (*M* = 21.0 years, *SD* = 2.8), middle-age adults (*M* = 47.0 years, *SD* = 6.0), and old adults (*M* = 72.0 years, *SD* = 7.9). *D* score calculation, exclusion criteria, and the analysis strategy were the same as in the prior studies (3.1% of participants who completed the MC-IAT were excluded from analyses).

Results

Sample sizes vary among the tests reported because of missing data. For participants in all six age groups, the order of implicit evaluations was the same. Each age group exhibited the most positive associations for children, followed by young adults, middle-age adults, and then older adults, all *t*s(1001–8328) > 2.4, all *p*s < .02, all *d*s > 0.05, average *d* = 0.27 (see Fig. 4 for results with age as a continuous variable).

Unlike the implicit evaluations of race and religion groups, the implicit evaluations of age groups did not reflect in-group favoritism. Nonetheless, there was evidence of a small in-group effect. When we restricted the sample to those participants whose ages ranged from 1 standard deviation below the estimated age for the young-adult images to 1 standard deviation above the estimated age for the middle-age images (i.e., ages

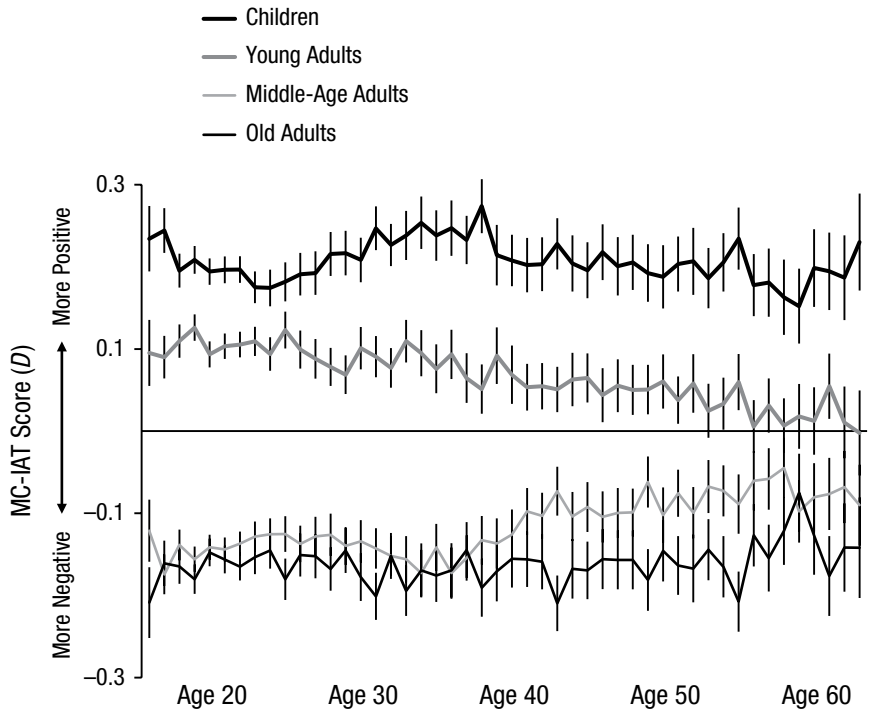

Fig. 4. Results from Study 3: Multicategory Implicit Association Test (MC-IAT) *D* scores for children, young adults, middle-age adults, and old adults as a function of participant's age (for all ages with more than 100 participants). Error bars denote 95% confidence intervals on the means.

18–53; n = 19,638), older age predicted more positive implicit evaluations of middle-age adults (r = .06, p < .001) and more negative implicit evaluations of young adults (r = –.08, p < .001). These effects were small enough that they did not disrupt the ordinal relations between any of the age groups. Table 4 presents descriptive statistics.

As in Studies 1 and 2, explicit evaluations did not show the same invariance as implicit evaluations. For example, although participants in their 20s, 30s, and 40s preferred children to older adults, the two groups were preferred roughly equally by participants in their 50s. Also, teenage participants and participants in their 20s preferred young adults to middle-age adults, whereas participants in their 30s and above evaluated young adults most negatively of all age groups. The Supplemental Material provides *t* and *d* values for implicit evaluations (Table S1), descriptive statistics and *t* and *d* values for explicit evaluations (Tables S2 and S3), descriptive statistics for each of the six implicit contrasts (Table S8), and correlations between implicit and explicit evaluations (Table S11).

Perceptions of likeability were strongly correlated with rated warmth of feelings (all *r*s > .54) and showed a

Table 4. Implicit Age Attitudes: Descriptive Statistics From Study 3

Participant's age group	Group evaluated											
	Children			Young adults			Middle-age adults			Old adults		
	n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
Teens	2,772	0.21	0.30	2,772	0.11	0.29	2,770	–0.15	0.27	2,769	–0.17	0.31
20s	8,362	0.19	0.30	8,362	0.10	0.28	8,369	–0.13	0.28	8,374	–0.16	0.31
30s	3,864	0.24	0.31	3,862	0.09	0.29	3,862	–0.15	0.28	3,865	–0.18	0.32
40s	3,306	0.21	0.31	3,299	0.06	0.28	3,294	–0.10	0.27	3,295	–0.17	0.33
50s	2,572	0.19	0.32	2,580	0.03	0.27	2,579	–0.08	0.28	2,578	–0.15	0.32
60s	1,008	0.20	0.32	1,008	0.01	0.27	1,008	–0.08	0.28	1,007	–0.13	0.33

Note: The table presents mean aggregate *D* scores (Greenwald, Nosek, & Banaji, 2003) from the Multicategory Implicit Association Test; more positive values indicate greater preference for members of the target group.

similar lack of invariance across the participant age groups. Perceptions of competence were more weakly correlated with likeability (all *r*s > .33) and warmth (all *r*s > .27). Also, all six age groups of participants showed the same ordinal pattern of competence ratings (middle-age adults > old adults > young adults > children), but this pattern was nearly the reverse of the hierarchy observed for implicit evaluations (for details, see Table S10 in the Supplemental Material).

General Discussion

Across three social domains, we found evidence for clear rules of implicit social evaluation that are largely invariant across racial, religious, and age groups. Rules of explicit social evaluation were less consistent. For race, the hierarchy of implicit evaluations placed the in-group at the top, followed by Whites, Asians, Blacks, and then Hispanics. For religion, the hierarchy of implicit evaluations again placed participants' own religion at the top, followed by Christianity, Judaism, Buddhism or Hinduism, and then Islam. For age, implicit evaluations placed children highest, followed by young adults, middle-age adults, and then older adults. These hierarchies of social evaluation complement evidence for distinct, consensual hierarchies for social power.

The results suggest that hierarchies of social identities are partly dependent on culturewide social structures and are pervasively embedded in social minds (Nosek & Hansen, 2008), particularly in the case of implicit evaluations, which are not endorsed and may even be contrary to conscious beliefs and values (Nosek et al., 2012). One interpretation of the difference between explicit and implicit measures is that implicit evaluations reflect the accumulation of experience, whereas explicit evaluations are qualified by idiosyncratic beliefs and values that are consciously decided and endorsed (Gawronski & Bodenhausen, 2006; Nosek & Hansen, 2008). Finally, the results suggest that implicit evaluative hierarchies reflect both in-group favoritism (Tajfel, 1978) and system justification (Jost et al., 2004).

Boundaries of in-group favoritism

In-group favoritism is seen in most social groups, particularly explicitly (Mullen, Brown, & Smith, 1992), but is not always seen in implicit evaluations (Jost et al., 2004; Rudman & Goodwin, 2004). The most dramatic lack of in-group favoritism that we observed was in the implicit evaluations of age groups in Study 3. An influence of in-group identity was swamped by the general evaluation that younger is better—even among the oldest participants. Why implicit in-group favoritism occurred for racial and religious but not age groups cannot be inferred directly from these data. One plausible explanation is that age categories are ambiguous and people can avoid identifying themselves as belonging to older age categories (e.g., Kleinspehn-Ammerlahn, Kotter-Gruhn, & Smith, 2008). This may have played a role, but the fact that the rank ordering from children at the top to old adults at the bottom remained constant even among people in their 60s suggests that this explanation is incomplete.

Another possibility is that the faces used to represent each age group evoked differences in attractiveness or caregiving responses in different age groups of participants. However, a similar pattern of implicit favoritism for young people over old people among participants across a wide age span has been observed with implicit measures using names instead of faces (Nosek et al., 2002).

Finally, research on in-group favoritism suggests a variety of boundary conditions for the effect. For example, high levels of belief in social dominance (Overbeck, Jost, Mosso, & Flizik, 2004) and high levels of perceived out-group negativity (Livingston, 2002) are associated with low levels of in-group favoritism. However, it is not clear that these boundary conditions are sufficient to account for the fact that in-group favoritism was found for racial and religious but not age groups.

Social status as a function of competence and warmth

Existing evidence shows substantial consensus for hierarchies of social power (e.g., Berry & Kalin, 1979; Hagendoorn et al., 1998; Kahn et al., 2009). The present results extend the evidence for pervasive hierarchies to implicit social evaluation—who is *good*. It is already well known that social evaluation and social power are not equivalent (e.g., Eagly, Mladinic, & Otto, 1991). In the present studies, members of each racial and religious group implicitly evaluated their own group atop the hierarchy, on average. Unless there is widespread misperception of group power among members of these groups, this finding suggests that social power is not the only contributor to implicit hierarchies. Moreover, in the age hierarchy, children were implicitly evaluated most positively despite having relatively little social power and the lowest level of perceived competence.

The present evidence adds to the existing literature by suggesting that both social power and social evaluation contribute to understanding social status, and that both may be embedded in implicit social cognition. A notable distinction is that social evaluation appears, in some cases, to be more sensitive to group membership than perceived social power is. Some groups perceive themselves to be the most good even if they recognize that they are not the most powerful.

From our perspective, the most promising means of understanding the interplay of power and evaluative hierarchies is to consider two dimensions of social

evaluation: competence and warmth (Fiske, Cuddy, Glick, & Xu, 2002). In Study 3, we measured competence and warmth explicitly but not implicitly. The data revealed distinct hierarchies of evaluation, with warmth ratings being closer to the implicit evaluations. However, documenting that perceived power is a function of warmth and competence hierarchies will require implicit measurement of both to estimate their independent and joint contribution.

Hispanics, not Blacks, occupy the bottom of the implicit racial hierarchy

A noteworthy side finding from Study 1 was that Black people generally received more positive implicit evaluations than Hispanic people. Past research has indicated that Blacks occupy the lowest rung of the racial status hierarchy (Sidanius, Levin, Liu, & Pratto, 2000). Recent work suggests that Hispanics may in fact occupy a position of lower status in the United States. For example, Hispanic men and women have lower weekly earnings than their White, Asian, and Black counterparts (U.S. Department of Labor, 2013). Our Study 1 now reveals that Hispanics are evaluated less positively on average than Blacks, at least implicitly.

Limitations of these studies

It is important to note that although our samples were extremely large, they are not representative of any definable population. It is possible that representative samples of the U.S. population would not show the same invariance in social evaluation of racial, religious, and age categories, though we cannot identify a plausible reason to expect this lack of generalizability. Further, it will be useful to extend this line of research to other forms of implicit measurement, and to other social domains, to document the extent to which implicit evaluations demonstrate invariance in social ranking in other areas.

We have focused on the rules of social evaluation in the aggregate, but the data revealed many unique effects that have potential research implications. For example, it remains unclear why Buddhism received such positive implicit evaluations from nonreligious participants and participants who did not belong to the target religions, or why Black participants were the only group to have a comparatively negative evaluation of White people in Study 1 (see Fig. 1). Furthermore, a reviewer noted that the structure of the MC-IAT may lend itself to alternative analysis strategies (e.g., dyadic analysis; Kenny & La Voie, 1984) that could provide additional insights into the structure of social evaluation. In short, the data set is larger and richer than the results we have presented here indicate. To facilitate additional research, we have made all data and materials available at the Open Science Framework (https://osf.io/zg2su/).

Conclusion

The present studies document invariance in implicit social evaluation across racial, religious, and age groups in the United States. These implicit evaluations appear to be dually sensitive to in-group identity and the relative status of other groups. That is, the rules of implicit social evaluation cannot be determined by in-group identity alone. An obvious next step is to clarify the origins of such implicit hierarchies, as well as their consequences for social judgment.

Author Contributions

J. R. Axt and B. A. Nosek developed the concept of these studies. J. R. Axt analyzed the data from Studies 1 and 3. C. R. Ebersole analyzed the data from Study 2. J. R. Axt drafted the manuscript, and B. A. Nosek and C. R. Ebersole edited it. All authors approved the final version of the manuscript for submission.

Acknowledgments

We thank Grant Garland for assisting with collecting and analyzing the data for Study 1, as well as Nicole Lindner for providing the data for Study 3.

Declaration of Conflicting Interests

B. A. Nosek is an officer and J. R. Axt is a consultant of Project Implicit, Inc., a nonprofit organization with the mission of "develop[ing] and deliver[ing] methods for investigating and applying phenomena of implicit social cognition, including especially phenomena of implicit bias based on age, race, gender or other factors." The authors declared that they had no other potential conflicts of interest with respect to their authorship or the publication of this article.

Supplemental Material

Additional supporting information can be found at http://pss.sagepub.com/content/by/supplemental-data

Open Practices

All data and materials have been made publicly available via Open Science Framework and can be accessed at https://osf.io/zg2su/files/. The complete Open Practices Disclosure for this article can be found at http://pss.sagepub.com/content/by/supplemental-data. This article has received badges for Open Data and Open Materials. More information about the Open Practices badges can be found at https://osf.io/tvyxz/wiki/view/ and http://pss.sagepub.com/content/25/1/3.full.

Notes

1. Study 1 had 105,293 started sessions; 97,641 participants provided data, and 60,611 completed the study (62% completion rate).
2. Pretesting revealed that adding prototypical surnames increased accuracy particularly for Hispanic faces.
3. Study 2 had 366,629 started sessions; 353,048 participants provided data, and 224,648 completed the study (64% completion rate).
4. This study had 54,665 started sessions; 49,014 participants provided data, and 29,982 completed the study (61% completion rate).

References

Berry, J. W., & Kalin, R. (1979). Reciprocity of inter-ethnic attitudes in a multicultural society. *International Journal of Intercultural Relations*, *3*, 99–112.

Bettencourt, B. A., Dorr, N., Charlton, K., & Hume, D. L. (2001). Status differences and in-group bias: A meta-analytic examination of the effects of status stability, status legitimacy, and group permeability. *Psychological Bulletin*, *127*, 520–542.

Campbell, K. E., Marsden, P. V., & Hurlbert, J. S. (1986). Social resources and socioeconomic status. *Social Networks*, *8*, 97–117.

Eagly, A. H., Mladinic, A., & Otto, S. (1991). Are women evaluated more favorably than men? An analysis of attitudes, beliefs and emotions. *Psychology of Women Quarterly*, *15*, 203–216.

Ebner, N. C. (2008). Age of face matters: Age-group differences in ratings of young and old faces. *Behavior Research Methods*, *40*, 130–136.

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*, 878–902.

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin*, *132*, 692–731.

Graham, J., Hawkins, C. B., & Nosek, B. A. (2012). [Ideology 2.0 data set]. Unpublished raw data.

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*, 197–216.

Hagendoorn, L., Drogendijk, R., Tumanov, S., & Hraba, J. (1998). Inter-ethnic preferences and ethnic hierarchies in the former Soviet Union. *International Journal of Intercultural Relations*, *22*, 483–503.

Hagendoorn, L., & Hraba, J. (1987). Social distance towards Holland's minorities: Discrimination against and among ethnic groups. *Ethnic and Racial Studies*, *10*, 120–133.

Hraba, J., Hagendoorn, L., & Hagendoorn, R. (1989). Foreign, different, deviant, seclusive and working class: Anchors to an ethnic hierarchy in the Netherlands. *Ethnic and Racial Studies*, *12*, 441–468.

Jost, J. T., Banaji, M. R., & Nosek, B. A. (2004). A decade of system justification theory: Accumulated evidence of conscious and unconscious bolstering of the status quo. *Political Psychology*, *25*, 881–919.

Kahn, K. B., Ho, A. K., Sidanius, J., & Pratto, F. (2009). The space between us and them: Perceptions of status differences. *Group Processes & Intergroup Relations*, *12*, 591–604.

Kenny, D. A., & La Voie, L. (1984). The social relations model. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 18, pp. 141–182). Orlando, FL: Academic Press.

Kessler, R. C., Mickelson, K. D., & Williams, D. (1999). The prevalence, distribution, and mental health correlates of perceived discrimination in the United States. *Journal of Health and Social Behavior*, *40*, 208–230.

Kleinspehn-Ammerlahn, A., Kotter-Gruhn, D., & Smith, J. (2008). Self-perceptions of aging: Do subjective age and satisfaction with aging change during old age? *Journals of Gerontology Series B: Psychological Sciences & Social Sciences*, *63*, P377–P385.

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition & Emotion*, *24*, 1377–1388.

Lin, N., Ensel, W. M., & Vaughn, J. C. (1981). Social resources and strength of ties: Structural factors in occupational status attainment. *American Sociological Review*, *46*, 393–405.

Livingston, R. W. (2002). The role of perceived negativity in the moderation of African Americans' implicit and explicit racial attitudes. *Journal of Experimental Social Psychology*, *38*, 405–413.

Minear, M., & Park, D. C. (2004). A lifespan database of adult facial stimuli. *Behavior Research Methods, Instruments, & Computers*, *36*, 630–633.

Mullen, B., Brown, R., & Smith, C. (1992). Ingroup bias as a function of salience, relevance, and status: An integration. *European Journal of Social Psychology*, *22*, 103–122.

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Harvesting implicit group attitudes and beliefs from a demonstration website. *Group Dynamics*, *6*, 101–115.

Nosek, B. A., Bar-Anan, Y., Sriram, N., Axt, J. R., & Greenwald, A. G. (2014). *Understanding and using the Brief Implicit Association Test: I. Recommended scoring procedures*. Unpublished manuscript, University of Virginia, Charlottesville, VA.

Nosek, B. A., & Hansen, J. J. (2008). The associations in our heads belong to us: Searching for attitudes and knowledge in implicit evaluation. *Cognition & Emotion*, *22*, 553–594.

Nosek, B. A., Hawkins, C. B., & Frazier, R. S. (2011). Implicit social cognition: From measures to mechanisms. *Trends in Cognitive Sciences*, *15*, 152–159.

Nosek, B. A., Hawkins, C. B., & Frazier, R. S. (2012). Implicit social cognition. In S. Fiske & C. N. Macrae (Eds.), *Handbook of social cognition* (pp. 31–53). New York, NY: Sage.

Nosek, B. A., Smyth, F. L., Hansen, J. J., Devos, T., Lindner, N. M., Ranganath, K. A., . . . Banaji, M. R. (2007). Pervasiveness and correlates of implicit attitudes and stereotypes. *European Review of Social Psychology*, *18*, 36–88.

Overbeck, J. R., Jost, J. T., Mosso, C. O., & Flizik, A. (2004). Resistant vs. acquiescent responses to ingroup inferiority as a function of social dominance orientation in the USA and Italy. *Group Processes & Intergroup Relations*, *7*, 35–54.

Rudman, L. A., & Goodwin, S. A. (2004). Gender differences in automatic ingroup bias: Why do women like women more than men like men? *Journal of Personality and Social Psychology*, *87*, 494–509.

Sidanius, J., Levin, S., Liu, J. H., & Pratto, F. (2000). Social dominance orientation and the political psychology of gender: An extension and cross-cultural replication. *European Journal of Social Psychology*, *30*, 41–67.

Sidanius, J., & Pratto, F. (1999). *Social dominance: An intergroup theory of social hierarchy and oppression*. New York, NY: Cambridge University Press.

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, *75*, 417–453.

Sriram, N., & Greenwald, A. G. (2009). The Brief Implicit Association Test. *Experimental Psychology*, *56*, 283–294.

Tajfel, H. (Ed.). (1978). *Differentiation between social groups: Studies in the social psychology of intergroup relations*. London, England: Academic Press.

Tajfel, H. (1982). Social psychology of intergroup relations. *Annual Review of Psychology*, *33*, 1–39.

Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. G. Austin (Eds.), *The psychology of intergroup relations* (pp. 7–24). Chicago, IL: Nelson-Hall.

U.S. Department of Labor, Bureau of Labor Statistics. (2013). *Labor force statistics from the Current Population Survey*. Retrieved from http://www.bls.gov/cps/cpsaat37.htm

Verkuyten, M. J. A. M., Hagendoorn, L., & Masson, K. (1996). The ethnic hierarchy among minority and majority youth in The Netherlands. *Journal of Applied Social Psychology*, *26*, 1104–1118.

Williams, D. R., Yu, Y., Jackson, J. S., & Anderson, N. B. (1997). Racial differences in physical and mental health: Socio-economic status, stress and discrimination. *Journal of Health Psychology*, *2*, 335–351.