

# Expanding Discrimination Research: Beyond Ethnicity and to the Web\*

Timur Kuran, *University of Southern California*

Edward J. McCaffery, *University of Southern California, California  
Institute of Technology*

*Objective.* This article aims to expand research about perceptions of discrimination both substantively and methodologically beyond the domains of race and ethnicity, relying partly on web-based surveys. *Methods.* We conducted parallel surveys over the telephone and the World-Wide Web, using standard random-digit dial (RDD) techniques for the former, and a large volunteer panel for the latter. *Results.* Both modes, phone and web, revealed that respondents consider discrimination based on physical appearance and economic status to be more prevalent than discrimination based on ethnicity. Respondents also reported that they themselves have been victimized more by physical appearance and economic-status discrimination than by ethnic discrimination. Significant differences emerged between the phone and web respondent pools, even after controlling for such independent variables as age, race, education level, and gender. *Conclusions.* People perceive discrimination across many aspects of social life, and appear more willing to reveal knowledge about controversial social phenomena on the web than on the phone.

There is a large scholarly literature spread across the social sciences on ethnic, especially racial, discrimination (Thernstrom and Thernstrom, 1997; Sniderman and Carmines, 1997). Researchers have paid less attention to other forms of discrimination, such as that based on economic status or physical appearance.<sup>1</sup> The relative lack of studies on physical-appearance

\*Direct correspondence to Edward J. McCaffery, USC Law School, Los Angeles, CA 90089-0071 (emccaffe@law.usc.edu). McCaffery will share all data and coding procedures with anyone wishing to replicate this study. The authors thank the USC-Caltech Center for the Study of Law & Politics for financial assistance; Carla VanBeselaere and especially Mike Alvarez for technical and scholarly advice; Kevin McCabe for useful discussions; and Yongchuan (Kevin) Bao, Stephanie Little, Sripad Motiram, Candice Mullings, Christina Yang, and especially Sung Han Tak for research assistance.

<sup>1</sup>The few scholarly articles found by searching major bibliographic databases are exceptions that prove the rule. Hamermesh and Biddle (1994) and Biddle and Hamermesh (1998), two important analyses of discrimination in favor of “good-looking” people, were published in high-impact economics journals, according to the classification of the ISI Web of Knowledge: the *American Economic Review*, ranked fifth in terms of impact among the 165 economics journals, and the *Journal of Labor Economics*, ranked 38th. As of September 2003, the first article had been cited 39 times in the thousands of ISI-covered journals, the second 10 times (<http://isi5.newsisknowledge.com>). A search in the Westlaw database, which contains

discrimination, dubbed “lookism” in popular discourse (Safire, 2000; Passell, 1996; Swisher, 1994; Trochek, 1990) contrasts with the persistent popular attention given to matters of beauty, supported by recent neurobiological research (Aharon et al., 2001). The shortage of research on economic-status discrimination contrasts with evidence of great economic inequalities (Phillips, 2002; Frank and Cook, 1995). These patterns raise questions of whether the priorities of researchers are in line with the perceptions of subjects. The substantive goal of this article is to expand the domain of discrimination research.

We also have a methodological goal: to expand the domain of tools available for discrimination researchers. The World-Wide Web is seeing increasing use among opinion researchers (Krosnick and Chang, 2000), a trend likely to continue as the biases of traditional phone surveys come under heightened scrutiny (Holbrook, Green, and Krosnick, 2003; Aquilino, 1994; Harwood and Leung, 2002). As web-based surveys gain visibility as barometers of opinion, however, questions of possible biases arise.

Most recent studies on *mode effects* in opinion surveys have contrasted telephone with face-to-face or self-administered interviews (Holbrook, Green, and Krosnick, 2003; Aquilino, 1994; but see Dillman et al., 2003). The central finding of this literature has been that people are relatively more guarded—less likely to reveal potentially controversial information—over the phone than in self-administered studies (Holbrook, Green, and Krosnick, 2003; Kuran, 1995; Aquilino, 1994). There is some disagreement about the effect of an actual interviewer on this *social desirability response bias*: the tendency for subjects to distort information or preferences, even to lie, so as to cultivate a favorable image in the eyes of the interviewer (compare Holbrook, Green, and Krosnick, 2003 and Kuran, 1995 with Dillman et al., 2003).

Methodological inquiries into web-based public opinion research, in contrast, have mainly focused on selection biases rooted in the “digital divide”—variations in computer use and literacy that make samples composed of web users unrepresentative of society as a whole (Couper, 2000; Nathan, 2001). The associated survey response bias is proving difficult to overcome. Researchers have shown that it is insufficient simply to weight subpopulations differentially to approximate their shares within the broader population of interest (Alvarez and VanBeseleare, forthcoming; Krosnick and Chang, 2000; Berrens et al., 2003). There may be unobserved demographic variations within the subpopulations of interest: the attitudes of college-educated white male web users, for example, may differ from those of college-educated white male nonweb users. Even controlling for demographic variables, web-based polling may differ from alternative modes (VanBeseleare, 2002). Insofar as the same individuals answer questions

differently over the web than over the phone, the data will be driven by the choice of survey mode as opposed to personal traits (Alwin and Krosnick, 1991).

We aimed to connect substance and method. Thus our study involves two large surveys designed to uncover beliefs about discrimination, one using the telephone and the other the web, on ethnic, physical-appearance, and economic-status discrimination. The telephone sample used traditional random-digit dial (RDD) techniques; the web sample used nonprobability, broad-based recruitment methods with mixed entry portals (Alvarez, Sherman, and VanBeselaere, 2003), what Couper (2000), in an analysis of web survey techniques, refers to as “volunteer panels.” We asked both respondent pools identical sets of questions about their perceptions of the prevalence of, and their personal experiences with, three distinct forms of discrimination. The conjunction of our substantive and methodological foci—respectively, three forms of discrimination and two modes of surveys—led to interesting social, political, and economic findings. Most important, our data suggest that concerns about discrimination on the basis of physical appearance and economic status are more prevalent than those about ethnic discrimination; that perceived social pressures are limiting the visibility of popular perceptions; and that the web holds promise in the field of discrimination research.

## **Survey Method and Hypotheses**

### ***Telephone Survey Pool and Selection Method***

The telephone survey consisted of 749 interviews conducted between January 13–22, 2002. Interviews were administered in both English and Spanish by the professional staff of a third-party provider, Interviewing Services of America. Respondents were selected using standard RDD techniques. Details of the survey design, response rate, descriptive statistics on respondents, and more are available from Alvarez et al. (2002a).

### ***Web Survey Pool and Selection Method***

The web survey consisted of 1,045 respondents. Potential survey participants were recruited using four methods: word of mouth; free web banner advertising on two separate banner sites; paid web banner advertising from third-party provider ValueClick; and direct subscription or co-registration with ValueClick. All respondents first provided information about themselves, including an email address; they were then emailed the survey’s URL and asked to participate. Participants were promised a chance at winning a \$50 gift certificate from an online bookseller. The survey was completed in

January 2002, co-extensive with the telephone survey described above. Details of the survey design, response rate, descriptive statistics on the respondents, and more are available from Alvarez et al. (2002b).

Note that we used a nonprobability web sample, acquired through mixed entry portals on a volunteer basis. Although such nonsystematic means of panel compilation raise the possibility of selection biases (Couper, 2000), they have been shown to produce statistically reliable results (Alvarez, Sherman, and VanBeselaere, 2003). Here, both the relatively large panel of respondents and the simultaneous use of a standard RDD phone technique as a check lend credence to the representativeness and reliability of our web data.

### *Questions*

We asked identical questions in the same order of both the web and the phone pools.<sup>2</sup> The questions fell into two parallel “clusters.” The three questions in each cluster dealt with ethnic, economic-status, and physical-appearance discrimination, respectively. The surveys asked about “discrimination based on ethnicity;” we did not use the more charged word “race” or any of its cognates.

Cluster A asked respondents about their judgments regarding the extent of discrimination in the United States today. We consistently asked respondents about their own perceptions (“in your judgment”), making no attempt to lay out the complexities or nuances of the meaning of “discrimination.” Our surveys captured respondents’ subjective beliefs pertaining to discrimination, according to their own understandings of the phenomenon. Specifically, the questions of Cluster A were:

1. In your judgment, how much discrimination based on ethnicity exists in the U.S. today?
2. In your judgment, how much discrimination based on a person’s economic status exists in the U.S. today?
3. In your judgment, how much discrimination based on a person’s weight, height, or physical appearance exists in the U.S. today?

The options given for each question were: A Great Deal; Some; Little; None; and Don’t Know.

<sup>2</sup>The survey described forms one component of a larger project in which different pools received different questions in different orders, to test for various framing, priming, and anchoring effects. In this particular component, the web and phone pools received identical questions in the same order. In all cases the questions considered here, 17–19 (Cluster A) and 20–22 (Cluster B), followed an open-ended question about the respondent’s ethnicity (“How do you describe your ethnicity to close friends?”). We include data from one-half of the phone survey (pickp1 in Alvarez et al., 2002a) and one-half of the web pool (Form A in Alvarez et al., 2002b). Conducting the analysis on pools roughly twice as large produced statistically identical results.

Cluster B inquired into respondents' personal experience (by self or family member) as a "victim" of discrimination. These questions, too, are subjective, as we did not define either "victim" or "discrimination." Specifically, they were:

4. Have you or anyone in your family been a victim of discrimination based on ethnicity?
5. Have you or anyone in your family been a victim of discrimination based on economic status?
6. Have you or anyone in your family been a victim of discrimination based on height, weight, or physical appearance?

Here the options were: Yes, Often; Yes, Sometimes; Never; and Don't Know.

### ***Hypotheses***

We believed that the greater anonymity and privacy of the web, compared to the telephone, would induce comfort with reporting infrequently investigated forms of discrimination. Hence we expected web respondents to reveal concerns over economic-status and physical-appearance discrimination more readily than phone respondents. We had no prior hypothesis regarding how these measures would compare with concerns over ethnic discrimination. On the basis of existing research (VanBeselaere, 2002; Krosnick, 1991), we believed that interviewer effects such as the social desirability bias would make respondents report relatively more ethnic discrimination over the phone. We also expected a strong positive correlation between the responses on Clusters A and B.

### **Results**

We first present the raw data in simple frequency distributions; the results are quite striking. Because of the differences in respondent pools, we then present the results of a multinomial logit (MNL) analysis, controlling for a variety of demographic variables as well as the mode—web or telephone.

#### ***Frequency Results***

Table 1 presents the frequency distributions for Cluster A, which concerned respondents' perceptions of discrimination.

On both the web and the phone, an overwhelming majority of respondents reported a belief that "a great deal" or "some" discrimination based on ethnicity, economic status, and physical appearance exists in the United States today. For all three categories, on both the web and the phone, over

TABLE 1  
Cluster A Frequencies

1. In your judgment, how much discrimination based on ethnicity exists in the US today?		
	Web	Phone
1. A great deal	39.9	35.5
2. Some	47.3	47.4
3. Little	10.1	10.8
4. None	1.5	2.4
5. Don't know	1.2	3.9
Observations	993	749
2. In your judgment, how much discrimination based on a person's economic status exists in the US today?		
	Web	Phone
1. A great deal	52.0	37.8
2. Some	38.9	44.2
3. Little	6.8	10.5
4. None	0.8	1.7
5. Don't know	1.5	5.7
Observations	992	749
3. In your judgment, how much discrimination based on a person's weight, height, or physical appearance exists in the US today?		
	Web	Phone
1. A great deal	56.5	43.4
2. Some	34.8	42.3
3. Little	6.5	9.2
4. None	1.4	1.6
5. Don't know	0.8	3.5
Observations	989	749
Pearson chi-square question		Web Versus Phone
1		17.03***
2		55.01***
3		40.55***

\*\*\* Means web and phone are statistically different at 1 percent level.

80 percent of respondents gave one of these two answers; over 90 percent did so for economic and physical-appearance discrimination. Pearson chi-square tests revealed, even in this simple setting, strongly significant differences between web and phone results. The most pronounced tendency was that web respondents were relatively more likely to answer "a great deal," and phone respondents to answer "some." Consistent with our hypothesis, the web-phone differences were greater on the questions about

economic status and physical appearance than on the more widely asked ones about ethnic discrimination. Respondents on the web were more likely to perceive “a great deal” of economic or appearance discrimination. This was true *both* compared to respondents on the phone (52 to 37.8 percent for economic status, and 56.5 to 43.4 for physical appearance) *and* compared to web respondents reporting ethnic discrimination (52 and 56.5 to 39.9 percent). Indeed, on both the web and the phone, the rank order of the “great deal” responses was the same, and surprising:

physical appearance > economic status > ethnic.

Table 2 presents the frequency distributions for Cluster B, the questions about personal experiences with discrimination.

Given the high number of respondents answering “a great deal” or “some” to each category of discrimination in Cluster A, what stands out is the low numbers reporting “often” or even “sometimes” to any of the categories, on either the phone or the web, in Cluster B. Most remarkably, a high percentage of the phone respondents reported that neither they nor their families had ever been victims of the designated form of discrimination. This was by far the dominant answer in regard to all three forms of discrimination (73.7 percent for ethnic discrimination, 72 percent for economic discrimination, and 69.8 percent for appearance discrimination). In each of these cases, the percentage was far in excess of the web respondents’ “never” answers (49.8, 32.8, and 33.4 percent, respectively). Indeed, there were rank-order reversals: on the web, but not on the phone, for both economic and physical-appearance discrimination, the “sometimes” answer was more common than the “never” answer. These differences between the web and phone respondents are highly significant according to Pearson chi-square tests. The many “never” answers on the phone contrast with the low percentages reporting “none” or even “little” to the questions in Cluster A.

### *Correlation Analysis*

Table 3 presents the Pearson correlation coefficients for the six questions in Clusters A and B, sorted by web and phone respondents. These coefficients are not especially high; none is above 0.5. This indicates that respondents differentiated among categories. The coefficients are all positive, and significantly different from 0, which means that a positive response is likely to be followed by another positive response.

Of particular interest is the shaded rectangle reflecting the relationship between Cluster A (Questions 1–3 in the columns) and B (Questions 4–6 in the rows). On both the web and the phone, the bold coefficients on the diagonal here are generally higher than those off diagonal. Thus, the relationship between the responses to Question 1 (about perceptions of ethnic discrimination) and Question 4 (about personal experience as a victim of

TABLE 2  
Cluster B Frequencies

4. Have you or anyone in your family been a victim of discrimination based on ethnicity?		
	Web	Phone
1. Yes, often	4.5	4.8
2. Yes, sometimes	30.4	19.9
3. Never	49.8	73.7
4. Don't know	15.2	1.6
Observations	991	749
5. Have you or anyone in your family been a victim of discrimination based on economic status?		
	Web	Phone
1. Yes, often	9.0	4.8
2. Yes, sometimes	44.7	20.6
3. Never	32.8	72.0
4. Don't know	13.5	2.7
Observations	988	749
6. Have you or anyone in your family been a victim of discrimination based on height, weight, or physical appearance?		
	Web	Phone
1. Yes, often	13.1	4.5
2. Yes, sometimes	37.6	22.3
3. Never	33.4	69.8
4. Don't know	15.9	3.3
Observations	990	749
Pearson chi-square question		
	Web Versus Phone	
4	143.11***	
5	270.90***	
6	244.36***	

\*\*\* Means web and phone are statistically different at 1 percent level.

ethnic discrimination) is tighter—the correlation coefficient is higher—than the relationship between Question 1 and either Question 5 or Question 6 (perceptions of ethnic discrimination vs. experience as victim of physical appearance or economic status). Likewise, the Question 1-Question 4 relationship is tighter than the relationship between Question 4 and either Question 2 or Question 3 (experience as a victim of ethnic discrimination vs. perceptions of appearance or economic discrimination). And so on.

On the web the Question 1-Question 4 correlation is higher than the off-diagonal correlations involving Question 1. On the phone, by contrast, the



TABLE 3  
Cluster A and B Correlations

Web Respondents						
Question	1	2	3	4	5	6
1	1.00					
2	0.39	1.00				
3	0.33	0.44	1.00			
4	<b>0.17</b>	<b>0.10</b>	<b>0.14</b>	1.00		
5	<b>0.13</b>	<b>0.32</b>	<b>0.21</b>	0.29	1.00	
6	<b>0.01<sup>x</sup></b>	<u>0.20</u>	<u>0.30</u>	0.27	0.39	1.00

  

Phone Respondents						
Question	1	2	3	4	5	6
1	1.00					
2	0.36	1.00				
3	0.37	0.40	1.00			
4	<b>0.18</b>	<b>0.17</b>	<b>0.15</b>	1.00		
5	<b>0.19</b>	<u>0.25</u>	<u>0.16</u>	0.37	1.00	
6	<b>0.19</b>	<u>0.17</u>	<u>0.23</u>	0.35	0.45	1.00

Pearson correlation coefficients. All significantly different from 0 at the  $p = 0.0000$  level, except for the coefficient marked <sup>x</sup>, which is different at the  $p < 0.002$  level.

Question 1-Question 5 and Question 1-Question 6 correlation coefficients are higher than the Question 1-Question 4 coefficient. Yet none of these differences is statistically significant, as measured by the Fisher  $r$ -to- $z$  transformation (Lowry, 2003). Consider now the correlations between Questions 2 and 5, and between Questions 3 and 6. These correlations, the underlined coefficients in Table 3, link perceptions of appearance and economic-status discrimination, respectively, with experiences as victims. Especially on the web, they are in general significantly higher than the off-diagonal pairs. For example, the Question 2-Question 5 correlation on the web (0.32) differs significantly from the Question 2-Question 4 correlation (0.10) at the  $z = 5.06$ , two-tailed  $p = 0$  level. It appears, then, that while experience as a victim of one form of discrimination makes one perceive that form as prevalent, especially for appearance and economic discrimination, generally it does not foster the belief that discrimination is also common on other dimensions.

Also significant is that, for the pairs of questions about less-studied forms of discrimination—those based on physical appearance and economic status—the coefficients are higher on the web than on the phone. Although the Question 1-Question 4 correlation on the web is statistically identical to that on the phone, the Question 2-Question 5 and Question 3-Question 6 correlations are higher on the web, with at least weak significance ( $z = 1.48$ , one-tailed  $p = 0.069$  for Question 2-Question 5;  $z = 1.58$ , one-tailed

$p = 0.057$  for Question 3–Question 6). Since personal experience with any given form of discrimination is likely to correlate with perceptions of its prevalence, this suggests that web respondents may have been more truthful or forthcoming about their own experiences as victims of discrimination.

### ***Multinomial Logit Analysis***

Because the respondent pools were different, in view of the nonsystematic web sample panel perhaps dramatically so, simple frequencies do not lend themselves to generalization. To isolate the effects of survey mode alone, as many variables as possible must be held constant. To this end, we conducted a MNL analysis. This method is appropriate because our dependent variable—the responses to Questions 1 through 6—involved a discrete choice, with limited options (Greene, 2003). MNL analysis generates predictions as to how changes in any given independent variable affect the discrete choice of the dependent variable.

We first tested several alternative models for goodness of fit, selecting one with five independent variables: age (coded in seven categories: under 18, 18–29, 30–39, 40–49, 50–59, 60–69, and over 70); gender dummy (1 = female); educational level (coded in six categories: did not complete high school, graduated from high school, attended college, vocational degree, graduated from college, graduate or professional degree); race dummy (1 = white, 0 = nonwhite); and mode dummy (1 = phone). Because the coefficients of MNL analysis are difficult to interpret, Table 4 presents the results of a first difference analysis, using Gary King's Clarify software (King, Tomz, and Wittenberg, 2003).

The figures in Table 4 provide a probability estimate of the likely change in each question (column) and response (row) category in response to moving a typical respondent from the web mode to the phone mode, holding all other variables (age, race, gender, educational level) constant. Thus, phone respondents were, *ceteris paribus*, 4.39 percent less likely than web respondents to give a 1 (a great deal) response to Question 1, which concerns the prevalence of ethnic discrimination; however, this result lacks statistical significance. Phone respondents were 13.2 percent less likely to answer "a great deal" to Question 2, which concerns physical-appearance discrimination; this result is statistically significantly different from 0. And so on.

Once again, the most striking mode effects are in Cluster B, Questions 4–6. Here, almost all the cells are statistically significant, and most of the magnitudes are dramatic. Shifting to the phone is likely to produce a significant reduction in the percentages answering "often" or "sometimes" (Responses 1 and 2) to these questions about personal experience with discrimination; and it would greatly increase the percentages likely to answer "never." For example, phone respondents, again *ceteris paribus*, are

TABLE 4  
Multinomial Logit Analysis, First Differences

Response	Question					
	1	2	3	4	5	6
1	-0.0439	-0.1318*	-0.0507*	-0.0055	-0.0440*	-0.0653*
2	0.0243	0.0654*	0.0533*	-0.1223*	-0.2087*	-0.1449*
3	-0.0061	0.0422*	0.0023	0.2899*	0.4067*	0.3985*
4	0.0085	0.0061	-0.0010	-0.1622*	-0.1539*	-0.1883*
5	0.0172*	0.0182*	-0.0039*	n.a.	n.a.	n.a.

\* Indicates differences significantly different from 0 at 95 percent confidence interval.

NOTE: First differences calculated using Clarify software (King, Tomz, and Wittenberg, 2003).

approximately 29, 41, and 40 percent more likely to answer “never” to the question of whether or not they or anyone in their families have ever faced discrimination on the basis of ethnicity, physical appearance, and economic status, respectively.

### Interpretation

The foregoing data furnish evidence of strong mode effects on answers to questions on discrimination. They also point to intriguing patterns concerning discrimination itself. Finally, there are interesting interactions between mode effects and relations among the alternative forms of discrimination.

Three findings stand out. One, from Cluster A, is that both web and phone respondents reported that in the United States there currently exists a significant amount of discrimination on the basis of physical appearance, economic status, and ethnicity, *and in that order*. This finding transcends mode effects. Two, also from Cluster A, is that web respondents were relatively more likely than phone respondents to report a great deal of physical-appearance or economic-status discrimination. MNL analysis that controlled for age, gender, race, and level of education confirmed this pattern. This finding at least weakly suggests that the social desirability bias is stronger on the phone than over the web: people apparently consider it relatively less “desirable” to report physical-appearance or economic-status discrimination than ethnic discrimination. Finally, from Cluster B, with respect to each form of discrimination, web respondents were much more likely to report that they or their families were among its victims. In each case, an overwhelming preponderance of phone respondents indicated that they themselves had “never” been harmed. This result, which survived MNL analysis, suggests also that the social desirability bias is greater on the phone

than on the web. Indeed, for our relatively less investigated domains of discrimination, correlations between personal experience and perception were higher on the web than on the phone. This, too, points to greater honesty on the web.

We discuss the mode effects and the substantive results in turn.

### *Mode Effects*

There is much evidence that survey mode can influence answers. On politically or socially sensitive issues, the “private preferences” that individuals register in anonymous surveys often differ dramatically from the “public preferences” they register in face-to-face interviews and focus-group discussions (Kuran, 1995). Public opinion researchers now call this effect the social desirability bias (Holbrook, Green, and Krosnick, 2003). Anything that gives respondents reason to fear retaliation for displaying an unpopular thought or preference can distort the results obtained. Katherine Bischooping and Howard Schuman (1992) showed that even the pen used by a face-to-face interviewer can influence poll results. Other studies show that on questions concerning controversial racial matters, the race of an interviewer can affect answers (Davis, 1997).

Our research gives strong reason to believe that the choice of survey mode—web or phone—will influence research on the sensitive topic of discrimination. Pinning down the sources of these mode effects is, however, more challenging than simply noting their significance. There are three broad, possible, and logically distinct sources of the effects.

*Selection Biases.* One reason for the observed differences between modes involves *selection biases*: the respondents differ. Unobserved variables may have influenced our respondent pools differently, especially on account of the differing means of compiling the pool, RDD versus volunteer panels (Couper, 2000).

There are reasons, however, to believe that these influences were limited.<sup>3</sup> The MNL analysis controlled for age, race, gender, and education level. The web-phone differences were also far greater on Cluster B than on Cluster A. If selection biases accounted for this discrepancy, the web would have attracted people who are *both* more likely than phone respondents to have experienced discrimination themselves (Cluster B) *and* roughly as likely to consider it socially significant (Cluster A), even with controls for gender, age, race, and education level. Although this cannot be ruled out, we believe it unlikely to account for the wide gaps within Cluster B responses.

Another possible source of selection bias comes into play during, rather than before, the survey completion process. We do not know how many

<sup>3</sup>Alvarez, Sherman, and VanBeselaere (2003) provide more general reasons why the influences are limited.

potential members of our web sample, if any, simply stopped answering once they arrived at our questions on discrimination and refused to submit the form. We know only that none of our phone respondents opted out midstream. Any web respondent who gave up midway would have left no trace in the remaining respondent pool. The consequent selection bias is different from item-nonresponse effects (VanBeselaere, 2002), which we could observe and that were not significantly greater on the web than on the phone.

It appears, however, that this within-survey selection bias, if indeed present, played no major role. It is on the web, rather than on the phone, that respondents were more likely to admit to being victimized by discrimination. For a within-survey selection bias to have explained the wide web-phone gap of Cluster B, web respondents answering “never” to the questions would have had to have dropped out in large numbers. But this bias, insofar as it arose, probably cut in the opposite direction: web *victims* might have dropped out on account of painful recollections, not because of their absence.

*Social Desirability Biases.* A second possible source of the observed web-phone difference is *social desirability bias*. This bias may manifest itself in terms of either *preference falsification*—the misrepresentation of preferences under the influence of perceived social pressures—or *knowledge falsification*—the misrepresentation of knowledge (or beliefs) to accommodate perceived social pressures (Kuran, 1995). What occurs here is knowledge falsification, because we asked subjects to report their perception of the prevalence of discrimination in society (Cluster A) and then to reveal their own personal experiences (Cluster B). Even in the absence of demographic differences, respondents may have differed across modes in terms of their honesty or truthfulness. For example, self-administered surveys, of which web-based surveys constitute the newest form, typically reveal more drug use than survey methods involving a live interviewer (Aquilino, 1994). Discrimination is a socially divisive issue, so there is reason to expect survey mode to matter. Web respondents may have felt relatively freer to share what they know about discrimination, as in traditional self-administered surveys (Dillman et al., 2003). Conversely, phone respondents may have been more likely to conceal their knowledge, perhaps out of privacy concerns or for fear of offending the interviewer (Holbrook, Green, and Krosnick, 2003).

Note that in Cluster A, the web-phone differences were more pronounced on economic and appearance discrimination than on ethnic discrimination. This is consistent with the closer attention that ethnic discrimination receives in the mass media, and hence with the perceived advantages of telling the phone surveyor what the respondent believes he or she might want to hear. Web answers may have suffered less than phone answers from the social desirability bias. In principle, the biases could have gone the other way, with web respondents falsifying more often. But the fact that phone

respondents were so highly likely to deny having experienced victimization on the Cluster B questions, and the tighter fit between Cluster A and B on the web than on the phone, suggests otherwise.

Both surveys respected standard conventions in regard to respondent privacy and anonymity. Why, then, might knowledge falsification have been less common on the web? One reason is that its technology affords more control of, and assurance about, privacy and anonymity. Respondents on the web were able to scroll down to see where they were heading. With little effort, they could check whether giving one answer rather than another would lead to unsettling follow-up questions (compare Dillman et al., 2003). They knew, in any event, that they could simply stop answering if the survey turned unpleasant and that they could do so without offending a live interviewer. In a phone survey, by contrast, respondents cannot know what lies in store. The mode affords no way to check whether answering “a great deal” to a question about experience with “discrimination based on physical appearance” will precipitate intrusive further questions about one’s weight, looks, or self-esteem. Concerns about privacy might explain in particular why a vastly greater share of the phone respondents answered “never” to the personal experience questions in Cluster B. Insofar as these questions were threatening, respondents could reduce the possibility of discomforting follow-up questions simply by denying that they themselves had been victimized.

Our results suggest that reporting discrimination on the basis of physical appearance or economic status is, indeed, considered relatively socially undesirable. Thus, in Cluster A, respondents appear to have presumed that phone surveyors believed, and expected to hear, that ethnic discrimination is common; these respondents would have been relatively reluctant to admit that physical-appearance discrimination is equally common or more so. In view of this possibility, it is all the more striking that, even on the phone, more respondents answered “a great deal” to questions about the extent of economic and appearance discrimination than to the question on ethnic discrimination.

*Transformation Effects.* A third possibility consists of *transformation effects*. Their source is not that the respondents are demographically different or give deliberately false answers to the surveyor. Rather, the truth itself is different, or appears to be different, on one survey mode than the other.

A simple factor involves the comfort levels that the media afford. Whereas a phone survey is typically conducted during the workday or early evening, and most likely in a rush, a web survey can be filled out at any time, when the respondent feels unrushed and relaxed (Holbrook, Green, and Krosnick, 2003; Dillman et al., 2003). A web subject need not feel trapped, as a phone subject might, when a surveyor calls; we used volunteer panels on the web, adding to this potential effect. Under relatively more relaxed conditions, web respondents may have found it easier to access their personal experiences.

Another type of transformation effect involves the psychophysical stimuli of the computer interface. Researchers have found that these differ across media (Birnbaum, 2001). Survey participants may respond differently on the web than on the phone simply because these media appeal to different senses: the sense of sight in one case and that of hearing in the other. Other physiological research shows that the prefrontal region of the brain is more active in interacting with humans than with computers (McCabe et al., 2001). These findings suggest that a web survey may trigger different sensitivities, evoke different memories, and give rise to different suppositions than a substantively identical phone survey. They give reason to believe that the choice of mode will affect the thought patterns determining how individuals respond to any given set of questions.

### ***Substantive Findings on Discrimination***

Both respondent pools considered appearance discrimination *more* significant than ethnic discrimination, with economic discrimination falling in between. This contrasts with the far greater media and social scientific attention received by ethnic discrimination. Because this result obtained on both large pools, it transcends mode effects.

Although this result surprised us, perhaps we should have expected it. A visit to any major bookstore or magazine stand, or a day in front of any television set, shows that Americans are intensely preoccupied with their appearance. Apparently, many Americans consider appearance a critical determinant of the treatments they receive. Neurobiological research (Aharon et al., 2001), a substantial literature in psychology (Low et al., 2003; Stice, Spangler, and Agras, 2001), and some economic analysis (Hamermesh and Biddle, 1994; Biddle and Hamermesh, 1998) support this sense that looks matter. Hamermesh and Biddle find that the “plainness penalty” for below-average looks is 5–10 percent of wages, and the “beauty premium” for above-average looks is almost as large. Unattractive women have lower labor participation rates, and they marry less-educated men. The impact of physical appearance is largely independent of occupation. With regard to economic-status discrimination, at a time of huge income and wealth gaps (Phillips, 2002), it is understandable that Americans are sensitive to class-based discrimination. In indicating that physical-appearance and economic-status discrimination might be more prevalent than ethnic discrimination, our respondent pools have pointed to something the media and academia have not adequately recognized.

### **Conclusions and Extensions**

It is clear that Americans consider discrimination a common phenomenon in the United States and, further, that they consider it based more often on

physical appearance or economic status than on ethnicity. Equally clear is that identifying the pertinent perceptions is a complex task. The phone may be less reliable than the web as a mode for uncovering sincere views about various forms of discrimination. Thus, our phone survey picked up the fact that physical-appearance and economic-status discrimination are considered common (Cluster A) but not the most obvious explanation for this fact—that people have personally experienced these forms of discrimination (Cluster B questions on personal experiences).

More research is needed to identify the precise nature, and then to quantify, the various mode effects that afflict opinion research on discrimination (see also Holbrook, Green, and Krosnick, 2003, in accord). Further findings may emerge from surveys that rule out specific effects. One can eliminate selection biases by having a given pool of subjects fill out identical surveys on the web and the phone, with standard controls for possible ordering effects (Nathan, 2001). This same technique would provide valuable insights into transformation effects and social desirability biases. One could also experiment with phone surveys that use recorded questions, or interactive voice response, in order to determine whether a live surveyor accentuates the social desirability bias; and with web surveys that have respondents hear recorded questions through an audio system in order to isolate the effects of reading questions on a screen (Dillman et al., 2003).

Most important, there are sound reasons to broaden the range of discriminatory domains studied. Our preliminary findings show that such expanded research may yield unexpected results. There may be other variables that Americans consider serious sources of discrimination. Possibilities include religion, national origins, age, marital status, sexual orientation, and educational achievement. It is time to broaden the agenda of discrimination research.

## REFERENCES

- Aharon, Itzhak, Nancy Etcoff, Dan Ariely, Christopher F. Chabris, Ethan O'Connor, and Hans C. Brieter. 2001. "Beautiful Faces Have Variable Reward Value: fMRI and Behavioral Evidence." *Neuron* 32:537–51.
- Alvarez, R. Michael, Robert P. Sherman, and Carla E. VanBeselaere. 2003. "Subject Acquisition for Web-Based Surveys: Political Analysis." *Political Analysis* 11:23–43.
- Alvarez, R. Michael, and Carla E. VanBeselaere. Forthcoming. "Web-Based Surveys." In *The Encyclopedia of Measurement*.
- Alvarez, R. Michael, Timur Kuran, Edward J. McCaffery, and Carla E. VanBeselaere. 2002a. *January 2002 Telephone Survey Codebook*. Available at <http://survey.caltech.edu/CodeJan2002Phone.pdf>.
- . 2002b. *Documentation for January 2002 Web Survey*. Available at <http://survey.caltech.edu/CodeJan2002Net.pdf>.
- Alwin, Duane F., and Jon A. Krosnick. 1991. "The Reliability of Survey Attitude Measurement: The Influence of Question and Respondent Characteristics." *Sociological Methods and Research* 20:139–81.



- Aquilino, William S. 1994. "Interview Mode Effects in Surveys of Drug and Alcohol Use: A Field Experiment." *Public Opinion Quarterly* 58:210–40.
- Berrens, Robert P., Alok K. Bohana, Hank Jenkins-Smith, Carol Silva, and David L. Weimer. 2003. "The Advent of Internet Surveys for Political Research: A Comparison of Telephone and Internet Samples." *Political Analysis* 11:1–22.
- Biddle, Jeff E., and Daniel S. Hamermesh. 1998. "Beauty, Productivity, and Discrimination: Lawyers' Looks and Lucre." *Journal of Labor Economics* 16:172–201.
- Birnbaum, Michael H. 2001. *Introduction to Behavioral Research on the Internet*. Upper Saddle River, NJ: Prentice Hall Press.
- Bischooping, Katherine, and Howard Schuman. 1992. "Pens and Polls in Nicaragua: An Analysis of the 1990 Preelection Surveys." *American Journal of Political Science* 36:331–50.
- Couper, Mick P. 2000. "Web Surveys: A Review of Issues and Approaches." *Public Opinion Quarterly* 64:464–94.
- Davis, Darren W. 1997. "The Direction of Race of Interviewer Effects Among African-Americans." *American Journal of Political Science* 41:309–32.
- Dillman, Don A., Glenn Phelps, Robert Tortora, Karen Swift, Julie Kohrell, and Jodi Berck. 2003. "Response Rate and Measurement Differences in Mixed Mode Surveys Using Mail, Telephone, Interactive Voice Response and the Internet." Presented at the Annual Conference of the American Association for Public Opinion Research. Montreal.
- Frank, Robert H., and Philip J. Cook. 1995. *The Winner-Take-All Society*. New York: Free Press.
- Greene, William H. 2003. *Econometric Analysis*, 5th ed. Upper Saddle River, NJ: Prentice Hall.
- Hamermesh, Daniel E., and Jeff E. Biddle. 1994. "Beauty and the Labor Market." *American Economic Review* 84:1174–94.
- Harwood, John, and Shirley Leung. 2002. "Hang-Ups: Why Some Pollsters Got it So Wrong this Election Day." *Wall Street Journal* November 8:a1.
- Holbrook, Allyson L., Melanie C. Green, and Jon A. Krosnick. 2003. "Mode Effects on Satisficing and Social Desirability." *Public Opinion Quarterly* 67:79–125.
- King, Gary, Michael Tomz, and Jason Wittenberg. 2003. *Clarify Software*, version 2.1. Available at (<http://gking.harvard.edu/stats/shtml>).
- Krosnick, Jon A. 1991. "Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys." *Applied Cognitive Psychology* 5:213–36.
- Krosnick, Jon A., and Lin Chiat Chang. 2000. *A Comparison of the Random Digit Dialing Telephone Survey Methodology with Web Survey Methodology as Implemented by Knowledge Networks and Harris Interactive*. Working Paper. Ohio State University.
- Kuran, Timur. 1995. *Private Truths, Public Lies: The Social Consequences of Preference Falsification*. Cambridge, MA: Harvard University Press.
- Low, Kathryn Graff, Swita Charanasomboon, Claire Brown, Gwen Hiltunen, Katherine Long, and Katherine Reinhalter. 2003. "Internalization of the Thin Ideal, Weight, and Body Image Concerns." *Social Behavior and Personality* 31:81–90.
- Lowry, Richard. 2003. "Significance of the Difference Between Two Correlation Coefficients." *VassarStats*. Available at (<http://faculty.vassar.edu/lowry/rdiff.html>).
- McCabe, Kevin, Daniel Houser, Lee Ryan, Vernon Smith, and Theodore Trouard. 2001. "A Functional Imaging Study of Cooperation in Two-Person Reciprocal Exchange." *Proceedings of the National Academy of Sciences* 98:1832–35.

Nathan, Gad. 2001. "Methodologies for Internet Surveys and Other Telesurveys." Presented at Proceedings of ETK 2001 International Exchange of Technology and Know-How and the Fourth New Techniques and Technologies for Statistics (NTTS) Seminar. Crete.

Passell, Peter. 1996. "A Study Shows Bias Against the Ugly: Should They Have Recourse?" *New York Times* January 11:d2.

Phillips, Kevin. 2002. *Wealth and Democracy: A Political History of the American Rich*. New York: Broadway Books.

Safire, William. 2000. "On Language." *New York Times Magazine* April 27.

Sniderman, Paul M., and Edward G. Carmines. 1997. *Reaching Beyond Race*. Cambridge, MA: Harvard University Press.

Stice, Eric, Diane Spangler, and W. Stewart Agras. 2001. "Exposure to Media-Portrayed Thin-Ideal Images Adversely Affects Vulnerable Girls: A Longitudinal Experiment." *Journal of Social and Clinical Psychology* 20:270–88.

Swisher, Karen. 1994. "Overweight Workers Battle Bias on the Job: Looks Discrimination Called Common, But Hard to Prove." *Washington Post* January 24:a1.

Thernstrom, Stephan, and Abigail Thernstrom. 1997. *America in Black and White: One Nation, Indivisible*. New York: Simon and Schuster.

Trochek, Kathy Hogan. 1990. "Obese Battle for Right to be as Fat as They Wish: End to Discrimination by Weight Seen as the Movement of the 90's." *Orange Country Register* November 12:e1.

VanBeselaere, Carla E. 2002. *Survey Response Quality: Shirking Behavior in Web and Telephone Surveys*. Working Paper. California Institute of Technology.