

NORC WORKING PAPER SERIES

The Effect of Survey Mode on Socially Undesirable Responses to Open-Ended Questions: A Mixed Method Approach

WP-2014-003 | OCTOBER, 2014

PRESENTED BY:

NORC at the University of Chicago

AUTHORS:

Danielle Wallace

E. C. Hedberg

Gabriel Cesar

Please send comments directly to principal author.

AUTHOR INFORMATION

Danielle Wallace¹

Arizona State University
School of Criminology and Criminal Justice
411 N. Central Ave.
Phoenix, AZ 85004
(602) 496-2349
danielle.wallace@asu.edu

E. C. Hedberg

NORC at the University of Chicago
and
Arizona State University
The Sanford School of Social and Family Dynamics
951 S. Cady Mall
P.O. Box 873701
Tempe, AZ 85287
ehedberg@asu.edu

Gabriel Cesar

Arizona State University
School of Criminology and Criminal Justice
411 N. Central Ave.
Phoenix, AZ 85004
gabrial.cesar@asu.edu

Word Count: 8,732

¹ Corresponding author.

Table of Contents

Abstract	1
Introduction	2
Social Desirability and Survey Mode	4
Web Surveys Generate Socially Undesirable Responses	4
Web Surveys Generate Fewer Socially Undesirable Responses	5
Study Context	6
Current Study	7
Methods	7
Data	7
Respondents	8
Stimulus	8
Coding Social Undesirability	9
Analysis Strategy	12
Results	13
Discussion	16
Conclusion	18
References	20
Notes	25

Table of Tables

Table 1: Descriptive Statistics and Probit Coefficients Predicting Likelihood of Online Survey	14
Table 2: Poisson and IRT Models Predicting Socially Undesirable Responses	15

Table of Figures

Figure 1: The Photographic Stimulus	9
--	---

Abstract

We model the chance of giving a socially undesirable response, one that violates social norms, by survey mode. We do this through a mixed method approach by evaluating open-ended responses to a photographic stimulus designed to generate undesirable responses. The study has 1,019 respondents; 55 percent took the survey online. We regress survey mode on social undesirability using a Poisson model predicting a respondents' number of socially undesirable responses and a Raush model predicting the likelihood of a socially undesirable response. We find that online surveys result in reduced rates of socially undesirable responses in the Poisson models, though survey mode is not significant in the Raush models. In sum, online surveys may not impact response content in open-ended questions.

Key Words: Online surveys, social desirability, open-ended questions, mixed methods

Introduction

With increased ease of survey creation through web-based survey engines (Couper and Miller 2008; Hargittai 2007; Kiesler and Sproull 1986), using web-based surveys to collect data is quickly becoming standard practice. Despite concerns about sampling, internet-based surveys have the same types of respondents generated from random digit dial (RDD) samples (Berrens et al. 2003). Berrens et al. (2003) show that with appropriate weighting, internet-based panel samples can look similar to RDD generated samples. As such, the use of online surveys is likely to grow, in addition to the increase seen from ease of administration (Couper and Miller 2008; Hargittai 2007; Kiesler and Sproull 1986). Thus, understanding how survey mode impacts social desirability in responses is both vital and timely for future survey research.

Web-based surveys come with some considerable advantages, chiefly that they are self-administered. Self-administered questionnaires (SAQs) increase the likelihood of eliciting unbiased information from respondents about sensitive topics such as experiences with discrimination, criminal behavior, or sexual activity (Kuran and McCaffery 2004, 2008). Of course, attention has been paid to socially desirable responses, or responses where the respondent is not truthful or accurate in order not to be seen in a negative light, in surveys and questionnaires (Holbrook, Green, and Krosnick 2003; Kuran 1995; Tourangeau and Smith 1996; Tourangeau and Yan 2007). However, perhaps the simplest and most used method is appropriate assurances of confidentiality and anonymity for the respondent. SAQs offer the ability to answer questions without face-to-face interactions. SAQs are a first line of defense when trying to obtain sensitive information from respondents by allowing them to express opinions or feelings they may not feel are socially desirable or that make them uncomfortable when discussing face to face. Thus, SAQs may be less likely to produce socially desirable responses on pro-social topics like church attendance (Kreuter, Presser, and Tourangeau 2008; Presser and Stinson 1998; Tourangeau and Yan 2007). However, there are situations where SAQs may be unable to reduce the possibility of biased, socially desirable responses when asking about sensitive topics.

Without the face-to-face cues individuals rely on to guide social interactions, online surveys give respondents the ability to answer survey questions without feeling the social presence of the researcher, which also enhances the sense of privacy individuals feel when taking surveys. Thus, people may respond in socially unacceptable ways (for this study, we call this social undesirability) that would typically not occur in a face-to-face survey setting. Further, evidence exists showing that the internet disinhibits individuals (Cassidy, Jackson, and Brown 2009; Herring et al. 2002; Phillips 2011) and there is reason to

suspect that this disinhibition may also find its way into survey responses (Christopherson 2007). In short, the lack of social contact coupled with feelings of enhanced privacy may lead to more candid (Kreuter, Presser, and Tourangeau 2008), and perhaps socially inappropriate, responses to survey questions.

On the other hand, web-based surveys may elicit more socially desirable responses from individuals, or at least, less informative responses. Because of the administration mode, online surveys may make subjects more leery of the confidentiality and anonymity that web surveys often suggest they offer (Fogel and Nehmad 2009; Young and Quan-Haase 2009) due to increased concern over privacy on the internet. As such, their responses might reflect levels of social desirability in an effort to control personal information.

To reconcile this paradox, we assess if taking a survey online influences the likelihood of giving a socially undesirable response through a mixed method approach. We explore this issue through a survey regarding neighborhood disorder, which is known as visual cues in a neighborhood's physical environment that convey the potential for issues of crime and safety (Skogan 1990; Wallace 2011). We present respondents with a photographic stimulus expected to generate socially undesirable responses from respondents. Because the survey is mixed mode—in-person and online—it provides an opportunity to measure whether the level of social undesirability differs by mode, and if it does, by what aspects of undesirability.

While there has been a wealth of research on social desirability in surveys, the vast majority of that work is quantitative. Scholars have a well-defined sense of when social desirability appears in surveys. Similarly, the moments in qualitative research where the researcher potentially influences data, such as during interviews or even through the researcher's presence, has also been well documented. However, to our knowledge, little work has addressed the presence of social desirability in open-ended questions. Here lies the contribution of our study. Mixed method approaches are excellent for adding insight to complex research problems (Branigan 2002; Creswell 2013; Stewart et al. 2008), and social desirability in research is one of these problems, given the increasing number of sensitive topics being broached in surveys (Tourangeau and Smith 1996; Tourangeau and Yan 2007). Without the mixed method approach of this study, or specifically examining open-ended answers to a photographic stimulus, the element of social undesirability found in this study would not be readily observable. This approach enabled respondents to express their perceptions of a neighborhood picture in their own words, and therefore, underlying themes of social undesirability emerged. In sum, this study capitalizes on one of the major benefits of mixed method research: allowing one method to interface with another for a more nuanced, powerful exploration of a common research dilemma such as social desirability (Mayoh and Onwuegbuzie 2013).

Social Desirability and Survey Mode

Survey mode may have a significant impact on how individuals respond to sensitive questions (for a review, see Tourangeau and Smith 1996); however, competing research exists which suggests that the desirability of individuals' responses to sensitive questions may be dependent on mode (Christopherson 2007). Of particular interest here is social *undesirability* in responses, or when responses to questions violate social norms that surround the context of the question. For example, consider a question that asks a respondent to describe the ideal racial composition of their neighborhood; a socially undesirable response to this question might include racial slang, stereotyping, and expressions of disgust. While the respondent could have expressed their view without their underlying attitudes being known, they chose not to. Below we discuss why online surveys may have an impact on the prevalence of socially undesirable responses.

Web Surveys Generate Socially Undesirable Responses

Scholarship shows that the online environment can be rife with bullying, name calling, and harassment (Cassidy, Jackson, and Brown 2009; Herring et al. 2002; Phillips 2011). Because of the enhanced sense of privacy people may feel when using the internet, individuals demonstrate little concern when discussing opinions, attitudes, or behaviors that in face-to-face situations would be unacceptable. As such, online surveys offer social anonymity, or “the perception of others and/or one’s self as unidentifiable because of a lack of cues to use to attribute an identity to that individual” (Christopherson 2007; p. 3040). Many studies have linked anonymity to anti-normative behaviors (Christopherson 2007) such as aggressiveness (Zimbardo 1969) and suicide (Mann 1981). Coffey and Woolworth (2004) describe a case where an online discussion board was established by a local newspaper as a forum to discuss a spree of possible hate crimes; they go on to compare those discussions to a local town hall meeting on the same topic. The authors found that the anonymous online forum contained overtly racist and violence-laden comments such as “IT’S TIME TO HANG THE PARENTS WITH THEIR CRIMINAL SPAWNS” (caps in the original; p. 6). Conversely, the town hall meeting was marked by a “noticeable lack of vengeful statements and...no threats of revenge or vigilante justice that had been so prevalent on the discussion board” (Christopherson 2007; p. 3049). Ignoring for a moment selection into attending a town hall meeting or responding to an online forum, it is plausible that online and in-person surveys work much the same way: respondents taking web-based surveys, when faced with stimuli or questions that provoke them to respond in a socially undesirable way, simply do so, where the socially undesirable tenor of their responses would not emerge in a face-to-face setting.

Coupled with the multitude of ways in which the web has an impact on social interactions, the web-based survey format may have impacts on responses. Particularly with open-ended responses, the online survey mode may present respondents with an opportunity to feel no pressure to respond in a pro-social way. Put simply, the anonymity often viewed as a strength of online surveys may have the unanticipated effect of fostering more extreme or socially undesirable responses than would face-to-face interviews or pencil and paper surveys.

Web Surveys Generate Fewer Socially Undesirable Responses

Recent news reports of governmental data collection serve to remind individuals that the internet is rarely anonymous (for example, see Savage, Wyatt, and Baker 2013). This would translate into fewer socially undesirable answers since respondents may feel increased monitoring when taking a web survey than when taking a paper and pencil survey.

Additionally, individuals are becoming savvier about privacy on the internet; for instance, scholarship demonstrates that there are growing concerns about privacy, particularly in regard to social networking websites like Facebook and MySpace (Fogel and Nehmad 2009; Young and Quan-Haase 2009). Survey takers may be more concerned about the disclosure of both personal information and their response to sensitive questions than individuals taking in-person paper surveys, regardless of whether the survey offers confidentiality or anonymity. Lastly, research on web-based surveys shows that respondents' perceived sensitivity of some items is more extreme when taking web-based surveys (Kreuter, Presser, and Tourangeau 2008), perhaps due to an increased perception that the internet is not private. Individuals' answers to sensitive questions may be less accurate and more socially desirable in an effort to control content and privacy.

Finally, some research suggests that respondents are more open regarding sensitive questions when taking surveys on the web (Kreuter, Presser, and Tourangeau 2008; Kuran and McCaffery 2004). When presenting questions regarding discrimination over the phone or online, Kuran and McCaffery (2004) found that people were more willing to discuss discrimination when taking an online survey. These findings reflect other studies demonstrating that the race of the interviewer can influence how people respond to questions regarding discrimination during an interview (Davis 1997). Similarly, Kreuter et al. (2008) found that individuals taking web surveys were more likely to report sensitive information when taking an online survey.

Of course, there is the possibility that people simply write down their opinions, attitudes, and feelings. While there is longstanding evidence that social desirability is an inherent problem in survey research

(Belson and Cleland 1986; Bishop, Tuchfarber, and Oldendick 1986; Sudman and Bradburn 1974; Tourangeau 2000), it may not be more or less a problem across different modes of survey administration. In other words, the prevalence of socially desirable or undesirable responses may not differ across administration mode or across different people within those modes.

Study Context

In order to assess differences in socially undesirable responses across survey mode, we employ a research design that uses mixed methods. Specifically, we use a photographic stimulus of neighborhood disorder that is expected to generate socially undesirable responses. Neighborhood disorder provides an excellent test case. In brief, disorder constitutes the visual cues within a neighborhood that signal to individuals that there is the potential for crime and personal victimization (Skogan 1990; Taylor 2001). Consequently, disorder is tightly linked with individuals' fear of crime: as individuals' perceive more disorder within a neighborhood, they tend to experience increased fear of crime and victimization (Taylor 2001).

Furthermore, there is strong evidence of socially desirable responses among individuals, particularly men, when they report fear of crime (Sutton and Farrall 2005). Sutton and Farrall (2005) argue that the underreporting of fear among men may help explain why men experience higher levels of victimization than women yet report much lower levels of fear and victimization risk. Given the link between fear of crime and disorder, the likelihood of generating socially undesirable responses from our stimulus is high.

Additionally, individuals may use disorder to cue other characteristics of the neighborhood, the most prominent being the ethnic and racial composition of a neighborhood. Sociological research has demonstrated that disadvantaged neighborhoods and the individuals living within them are often seen in an unfavorable light (Goffman 1963; Smith 1986; Wacquant 1993; Werthman and Piliavin 1967). This is due to the strong and durable link between race, poverty, crime, and racial segregation in the United States (Wacquant 1993). Consequently, many individuals may view neighborhood residents' skin color or the racial composition of a neighborhood as a primary indicator for the potential for crime and victimization. Furthermore, skin color is a particularly easy cue to use when individuals need to navigate their environment: "The human tendency to quickly categorize racial and other groups, our ability to easily observe skin color, and our sensitivity to the opinions of others in the form of reputations or identities that stigmatize areas, makes for a potent combination" (Sampson 2012; p. 125). Consequently, we expect there to be some stereotyping and racial animus present in the responses, even though no individuals are featured in the picture.¹

Current Study

This study examines whether the survey mode (i.e., in-person vs. online) had an impact on the likelihood of a socially undesirable response. We did this by presenting respondents with a photograph of neighborhood disorder that is expected to elicit strong, potentially biased responses. We then tested whether subjects' responses to a question in the online version of the survey were more or less likely to result in a socially undesirable response than those responses from the in-person, paper and pencil version of the survey. We explored three types of socially undesirable responses: *inappropriate descriptors* (e.g., “shitty,” “assholes”), *blatant stereotyping* (e.g., “Filled with Hispanics. Some of them gang members or trouble-making kids. Probably not too educated of people and do not respect their neighborhood. Filled with trouble—poorer class people.”), and *express distaste* (e.g., “Oh hell nooooo”). The coding protocol is discussed in greater detail below. We also used a variety score, or the number of undesirable codes triggered to examine how many types of undesirable responses subjects gave. For example, if an individual used inappropriate descriptors alone, this measure was coded as 1. If their response contained blatant stereotyping in addition to inappropriate descriptors, then this measure was coded as 2, and so on. We analyzed our data using two methods: (1) a simple count model (Poisson) to estimate the number of types of undesirable responses (the variety score) and (2) a multilevel item response model (Reise and Moore 2012) or Rasch model, where responses were nested within respondents. Specialty effects for type of undesirable response were estimated (Schreck et al. 2012) for both models. Given the potential for the stimulus to generate socially undesirable responses and the research showing that the online environment is conducive to violating social norms (Cassidy, Jackson, and Brown 2009; Herring et al. 2002; Phillips 2011), we expect that online respondents will be more likely to give socially undesirable responses than respondents who took the survey as a paper and pencil survey.

Methods

Data

Our data originated from the Perceptions of Neighborhood Disorder and Interpersonal Conflict Project (PNDICP), which was a study aimed at investigating how individuals perceive and interpret neighborhood contexts and interpersonal conflicts. Students from a large southwestern university in the United States were recruited to be a part of this project. The university from which the sample was drawn conducts a large number of classes online; consequently, two modes of administration were created, in-person and web-based, in an effort to sample both types of students. The total sample included 1,371

enrolled students.² The respondents totaled 1,056, yielding a response rate of 77 percent. The mode-specific response rates differed, with an 86 percent response rate for in-person survey administration and a 67 percent response rate for web-based survey administration. We calculated the response rate for each course surveyed using the number of students currently enrolled in the course as the denominator. Thus, our response rates did not distinguish between not attending the class on the day of the survey and refusal. The lowest response rate was for an online course at 42 percent of students responding, with the highest response rate corresponding to an in-person class at 98 percent of students responding. All instructors who participated in the administration of the survey were given the option to offer extra credit to their students for participation. The web survey was administered using a web survey construction tool, Qualtrics, which allows for complex randomization and contingency questions. Nearly 55 percent of respondents took the online survey.

Respondents

As noted before, our respondents were students from a large southwestern university. Nearly 83 percent of respondents reported being between the ages of 20 and 29, while 86 percent reported being single. Variation in race and ethnicity were more pronounced: 53 percent of the respondents were white, and about 30 percent of the respondents were Hispanic, while 7 percent were black. While a student sample lacks wide generalizability, the admission rate of the university where the study was conducted was approximately 89 percent; thus, the sample was heterogeneous enough to generate differences across respondents. The final sample size of open-ended responses to the stimuli photograph in the PNDICP was 1,019.

Stimulus

The stimulus for this study is a photograph of neighborhood disorder. Given that individuals' disorder perceptions do vary (Hipp 2010; Sampson and Raudenbush 2004; Wallace 2011), and that part of that variation can be attributable to racial characteristics such as race and poverty within neighborhoods (Latkin et al. 2009; Sampson and Raudenbush 2004; Taub, Taylor, and Dunham 1984), the photographic experiments³ in the PNDICP were designed to hold the neighborhood constant to inhibit bias coming from one's personal neighborhood context. Consequently, interpretations of the photograph are attributable to the individual and not to the neighborhood context. This enabled us to discern not just interpretive themes, but also issues of social undesirability.

Our stimulus, as seen in Figure 1, is a photograph of the west side of Chicago, in a neighborhood that is low to lower-middle class in income and is largely Hispanic.⁴ The photograph contained a religious mural, street art, and graffiti. The murals hinted at the racial composition of the neighborhood. Here, we tacitly cued for the racial composition of the neighborhood. Upon seeing the photograph, respondents were asked to give their immediate or gut reaction and finish the following statement: “Based on this photograph, I get the sense that this neighborhood is....” The open-ended response format of this question enabled respondents to reveal how they perceived the neighborhood in the photograph.

Figure 1: The Photographic Stimulus



Coding Social Undesirability

In order to code for social undesirability, a team of coders was assembled to identify and discuss common themes that emerged from the open-ended responses (Saldaña 2009). The coding team was assembled

with researchers who had immersed themselves in the data through the process of reading and entering the data, as well as reviewing the online responses in depth. Coding meetings involved an iterative process where each team member suggested themes identified as common while examining the data (Maxwell 2005; Saldaña 2009). Because the operationalization of social undesirability is potentially subjective (Crowne and Marlowe 1960; Edwards 1957), initial research meetings were focused on, first, identifying which responses were socially undesirable and, second, the emergent themes of social undesirability in the responses. This process generated three primary types of social undesirability: *inappropriate descriptors*, *blatant stereotyping*, and *express distaste*. Once these types of undesirability were established, MAXQDA 10.0 was used to organize the data and attach codes to relevant responses. All data was coded independently, and the results were compared for consistency and accuracy.

Responses were coded as socially undesirable when they appeared to violate some social norm surrounding the context of the question. Here, that context was answering the question “Based on this photograph, I get the sense that this neighborhood is...” in response to viewing the stimulus photograph in Figure 1. Thus, socially undesirable responses are those that violate what is seen as an acceptable way to describe or discuss a neighborhood. These responses tend to be culturally or racially insensitive, use slang or swearing, or demonstrate the respondent’s dislike of the neighborhood in the photograph. Approximately 20 percent of responses contained some element of social undesirability.

As noted above, socially undesirable responses were categorized into three types: inappropriate descriptors, blatant stereotyping, and express distaste. *Inappropriate descriptors* include responses containing profanity, such as “shit hole” and “gun-toting assholes,” as well as the use of “ghetto.” In addition, responses that contained culturally insensitive terms, like ghetto, were also coded as inappropriate descriptors.

There was extensive discussion among the coding team over the inclusion of “ghetto” in the inappropriate descriptors category. Given both the historical and contemporary cultural context of the term ghetto (Wacquant 1993; Werthman and Piliavin 1967), as well as the enduring effects of popular perceptions of ghetto-ness (Anderson 2012), we argue that the term ghetto is not a socially appropriate way of describing a place. Anderson (2012) best explains the idea of the ghetto:

“Today the ghetto is commonly regarded at best as the home of those black people who have been left behind by racial progress and, at worst, as a place inhabited by those who have failed to assume personal responsibility for themselves and their families” (p. 14).

In this sense, the idea of the ghetto is rife with assumptions regarding race, poverty, and personal responsibility. In light of Anderson's description of what it means for a person or group to be perceived as associated with the ghetto, we coded the use of this term as socially undesirable and include these responses in the inappropriate descriptors category.

The second type of social undesirability that emerged was the use of *blatant stereotyping*, or overtly linking certain types of people to the places in the photographs. Responses like "lazy people live here" and "dominated by guilt ridden Hispanic Catholics who have kids who enjoy tagging up the place" typify the presence of blatant stereotyping in the responses. These statements are coded as blatant stereotyping because respondents made strong assumptions about the characteristics of neighborhood residents. This coding practice is strengthened by the photograph itself: because the photographic stimulus does not portray individuals, any reference to people or residents in responses was derived only from the physical cues in the photograph. As with inappropriate descriptors, the coding of a text segment as blatant stereotyping was reserved for the most extreme cases.

Of particular interest in the current study was the racially based stereotyping by respondents, which is particularly relevant for the blatant stereotyping code. Given the religious, ethnic, and national components of the mural pictured in the stimulus, this was not unexpected. Specifically, the mural is composed of a Catholic saint (the Virgin Mary), depicted as a Hispanic woman (referred to as La Virgen de Guadalupe), who is supported by an angel whose wings are colored as the flag of the Republic of Mexico. Yet, even with the prevalence of race in the responses, some appeared to be impartial responses to cultural cues while others were clearly biased against certain ethnic and racial groups.⁵ The response "[O]verrun by blacks & Mexicans, very poor, poorly educated and hopeless" is an example. As such, only those responses where race/ethnicity was mentioned and there was clear bias were marked as socially undesirable and placed into the blatant stereotyping category.

Finally, *express distaste* emerged as the third theme of social undesirability. Unfiltered or expressions of distaste describe responses containing disapproval or a disinclination to participate in a community context. "OH, NO!" is a typical example of the use of express distaste in responses in the current sample; here, the words "oh no" not only express the respondent's desire to not interact with the type of community in the photograph, the capitalization and exclamation point also drive the tenor of the response.

As a note on the coding protocol, we did not rank the relative strength of the coding classifications. For instance, while “very ghetto” is technically stronger wording than simply “ghetto,” we focused on capturing only the socially undesirable nature of the response over the perceived severity of phrase.

Analysis Strategy

Our analysis strategy involved estimating the effect of survey mode on the likelihood of socially undesirable statements. We weighted the data by the inverse propensity weights due to the non-random assignment to the online or in-person survey mode.

For our first analysis, we used a simple Poisson model to predict the number of undesirable responses (0-3) each respondent made. That is, for the j^{th} respondent, we fit the simple model

$$\ln(y_j) = \beta_0 + \beta_1 MODE_j,$$

where β_l measures the effect of mode on the rate of undesirable response types. Mode was coded as -0.5 for in-person and 0.5 for online to render the intercept as an average.

The second model involved a Rasch model. To do this, we structured our data in “long” form so that each respondent received three rows of data. Each row corresponded to one of the three types of socially undesirable responses. For each row, we indicated whether that respondent was coded with a particular response with a dichotomous 1 for “yes” and 0 for “no.” That is, our outcome is whether the j^{th} respondent responded in a socially undesirable way for the i^{th} socially undesirable type y_{ij} , and our model is

$$\ln\left(\frac{\Pr(y_{ij} = 1)}{\Pr(y_{ij} = 0)}\right) = \beta_{0j} + \beta_{1j} GENERALIZATION_{ij} + \beta_{2j} DISTASTE_{ij} + r_{ij}$$

where

$$\beta_{0j} = \gamma_{00} + \gamma_{01} MODE_j + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} MODE_j$$

$$\beta_{2j} = \gamma_{20} + \gamma_{21} MODE_j$$

Again, *mode* was coded as -0.5 for in-person and 0.5 for online. This rendered the intercept the average likelihood of a socially undesirable response and not the likelihood of in-person surveys. We then created two other indicators to identify the undesirable response as either generalization or express distaste (inappropriate descriptors was the reference). Each of the indicators is deviance coded, whereby 1

indicated the specific undesirable response, -1 indicated the reference undesirable response, and 0 for all other undesirable responses. This coding scheme in effect preserved the intercept as the population average and not the likelihood of the reference undesirable response. Finally, we interacted mode with each of the undesirable response variables to estimate the moderating effect of the mode on the undesirable response type.

In this model, γ_{00} measured the log-odds of making an undesirable response, on average, and γ_{01} measured the main effect of mode on this likelihood. Beyond these effects, γ_{10} and γ_{20} measured the specialization effects of the undesirable response types of Generalization and Distaste relative to the population average, and γ_{11} and γ_{21} measured how this specialization differs for online surveys. Since our outcome is over dispersed, we include a level-1 residual, r_{ij} . Such a parameter in our model allows for over dispersion of our outcome due to the low base rates.

As mentioned above, whether the respondent took the online survey or the in-person survey was not randomly assigned, but was based on whether the respondent was taking either an online or in-person course. This required us to take into account the propensity of taking the survey online (i.e., treatment). To estimate the propensity weights, we fit a probit model predicting the likelihood of taking the survey online using the demographic variables available: Race, $\ln(\text{Age})$, an indicator showing the student lives locally, and gender. The location indicator was an important measure because some of the online students lived in a different region of the country. We then predicted the individual’s propensity to take an online survey and created weights equal to (Austin 2011a, 2011b):

$$w_i = \frac{X_i}{e_i} + \frac{1 - X_i}{1 - e_i}, X_i = \begin{cases} 1 & \text{if } i \text{ took survey online} \\ 0 & \text{otherwise} \end{cases}, e = \text{propensity} .$$

Results

Table 1 showcases the descriptive information for the dependent variables. The average number of types of undesirable responses was 0.20. About 15 percent of all respondents were coded as giving an undesirable response. Six percent of the respondents were coded with inappropriate descriptors, 10 percent used blatant stereotyping, and 4 percent were coded as expressing distaste. The descriptive information for the demographic variables and the probit coefficients of the demographic variables predicting online surveys are also shown in Table 1. We find that the “other” racial group—mostly Southeast Asians—was associated with in-person surveys. Older respondents were also more likely to

take online surveys. Local students were associated with in-person surveys, while females were associated with online surveys.

Table 1: Descriptive Statistics and Probit Coefficients Predicting Likelihood of Online Survey

Variable	Mean	Min	Max	Probit Coefficient Predicting Online Survey ^a
Socially Undesirable Response				
Number Undesirable Response Types	0.20	0	3	
Bad Adjective	0.06	0	1	
Generalization	0.10	0	1	
Distaste	0.04	0	1	
Online Survey	0.55	0	1	
Controls				
Race				
Black	0.02	0	1	0.30
Hispanic	0.24	0	1	-0.14
Other	0.06	0	1	-0.47**
Refuse/Don't Know	0.02	0	1	-0.12
Age	24.03	18	67	
ln(Age)	3.15	2.89	4.2	1.37***
Local Student	0.93	0	1	-0.49**
Sex				
Female	0.53	0	1	0.35***
Refuse/Don't Know	0.01	0	1	-0.06

Notes: N = 1,019 respondents; a: Intercept = -3.87, pseudo R2 = 0.06, ** p < 0.01, *** p < 0.001

We first fit a Poisson model to predict the rate of undesirable responses using the propensity weights. We found that online surveys reduced the rate of undesirable responses by 21 percent. The results are presented in Table 2. According to the model, the estimated number of undesirable responses for in-person surveys was about 0.21 or $exp(-1.663 - 0.5 * -0.228)$, and the estimated number of undesirable responses for online surveys was about 0.17 or $exp(-1.663 + 0.5 * -0.228)$.

Table 2: Poisson and IRT Models Predicting Socially Undesirable Responses

	Poisson Model ^d	IRT Model ^c	
		Unit Specific Model	Population Average Effects ^d
Fixed Effects			
Intercept	-1.663 *** (0.051)	-4.715 *** (0.124)	-2.767 *** (0.092)
Mode ^a	-0.228 * (0.102)	0.002 (0.249)	-0.123 (0.184)
Blatant Stereotyping ^b		0.957 *** (0.062)	0.515 *** (0.041)
Mode ^a × Blatant Stereotyping ^b		-0.917 *** (0.124)	-0.461 *** (0.082)
Express Distaste ^b		-0.874 *** (0.068)	-0.499 *** (0.051)
Mode ^a × Express Distaste ^b		0.400 ** (0.136)	0.209 * (0.102)
Random Effects			
Respondent Level VC	NA	6.576	NA
VPC=VC/(VC+π ² /3)	NA	0.667	NA
Log Likelihood Start	-1104.9724	-2132.187	
Log Likelihood Final	-1104.9724	-2056.483	
Pseudo R ²	0.002	0.036	

Notes: N = 1,019 respondents; a: Coded as -0.5 for in-person survey, 0.5 for online; b: Deviance coded to force intercept to be population average; c: Model fit as over-dispersed HGLM with logit-link; d: Population average effects do not assume specific random effects but describe plausible effects in the larger population; * p < 0.05, ** p < 0.01, *** p < 0.001

We then fit Item Response Theory (IRT) models to the long-formatted data. The results are presented in two ways, as “unit specific” results and as “population average” results. Unit specific effects hold the estimated variance components constant for interpretation, thus reflecting the *between respondent variability*, while the population average models are more consistent with traditional single-level models. The practical difference between the two methods is that unit specific models are designed to elaborate on the variance components but may not have fixed coefficients that reflect the averages present in the data. Population average models, which ignore the random effects, present coefficients that reflect the observed data. The reason we present both in this study is to show how much the between-person variance matters: there is a large difference between the population average and unit specific fixed effects.

While we did not observe main effects of the survey mode (the coefficient was essentially 0), we did find that mode moderated the specialization of social undesirability. In the unit specific results, we found that

the odds of blatant stereotyping for in-person surveys were over twice ($\exp(0.96) = 2.6$) the average, and that the odds of express distaste in in-person surveys were less than half ($\exp(-0.87) = 0.4$) those of average. However, the effect of generalization for online surveys was statistically 0 when we accounted for the survey mode interaction with generalization. A simple *t*-test combining the coefficients and their variances resulted in $t = 0.29$ or $(0.957 - 0.917) / \sqrt{(0.062^2 + 0.124^2)}$, which is not statistically significant. Thus, there was no generalization specialization for online surveys as there was for in-person surveys. The distaste specialization was reduced in online surveys as well and has a significant *t*-test of -2.94 or $(-0.847 + 0.400) / \sqrt{(0.068^2 + 0.136^2)}$. This means that odds of distaste in online surveys was about 36 percent less than average ($\exp(-0.447) = 0.64$).

We found similar effects in the population average coefficients, we again did not observe main effects of the survey mode (the effect is not statistically significant), and we still found that mode moderates the specialization of social undesirability. In the population average set of effects, we found that the odds of generalization for respondents who took in-person surveys were over 60 percent more ($\exp(0.52) = 1.68$) than that of average, and that the odds of distaste in in-person surveys were 40 percent less than ($\exp(-0.50) = 0.6$) those of average. Again, the effect of generalization for online surveys was statistically 0 when we accounted for the mode interaction with generalization. A simple *t*-test combining the coefficients and their variances resulted in $(0.525 - 0.461) / \sqrt{(0.041^2 + 0.082^2)} = 0.70$, which is not statistically significant. The distaste specialization was reduced in online surveys as well, but not eliminated with a *t*-test of $(-0.499 + 0.209) / \sqrt{(0.051^2 + 0.102^2)} = -2.54$. This means that the odds of distaste in online surveys were about 25 percent less than average ($\exp(-0.29) = 0.75$).

In conclusion, we found no evidence that online surveys increase the likelihood of undesirable response. However, when we examine the moderating effects of survey mode, we find that the type of undesirability is far less specialized in online surveys. This suggests that when there is social undesirability expressed in online surveys, the type of undesirability is harder to predict.

Discussion

Since the upsurge in the use of web-based surveys, survey researchers have been concerned with whether web-based surveys generate different responses than paper and pencil or face-to-face surveys. Related issues are social desirability or undesirability in responses, or when responses are inaccurate due to individuals changing their answer. Because of the unique social environment of the internet, where individuals can hide their identity and be dishonest without many consequences (de Leeuw 2012; Miller et al. 2002; Richman et al. 1999), understanding when social (un)desirability in surveys is a concern

(Christopherson 2007) is important. We address the issue of social undesirability across mode of administration by using a mixed method design involving qualitative coding and a photographic stimulus. Our primary findings are discussed below.

First, we found that online survey respondents were less likely to express socially undesirable responses to our stimuli than respondents who took the paper and pencil survey. Evidence from online communities and forums would suggest that more socially undesirable responses would come from the online survey respondents (Blevins and Holt 2009; Christopherson 2007; Hardaker 2010; Herring et al. 2002; Phillips 2011); however, we found the opposite. Thus, we concluded that for stimuli and questions aimed at eliciting a potentially sensitive and socially undesirable response, online surveys were unique in that respondents offered more socially acceptable responses.

We offer two possible explanations for this result. First, these results lend support to research suggesting that internet consumers are growing more aware of online privacy issues (Fogel and Nehmad 2009; Young and Quan-Haase 2009) and may be more likely to think that online surveys are not fully anonymous or confidential. Subsequently, it may be that respondents are more cautious when answering questions in web-based surveys. Second, this finding also supports work demonstrating that respondents have heightened concern over sensitive questions when fielded in an online survey (Kreuter, Presser, and Tourangeau 2008). Our findings that individuals who took the survey online were less likely to convey socially undesirable answers strongly suggests that the online format conveyed the sensitivity of items more readily to subjects. While web-based surveys enable the creation of cheap and easy surveys, a consequence of mode of administration may be that respondents are less trusting of our assurances of anonymity and confidentiality online. For bubble-sheet paper and pencil surveys, claims of privacy have face validity, whereas respondents may not perceive online surveys as being anonymous and confidential despite researchers' assurances.

The second and perhaps more interesting result stemmed from the differences in specialization (inappropriate descriptors, distaste, or generalization) within socially undesirable responses. Individuals who took in-person surveys were far more likely to offer blatant stereotyping and less likely to offer distaste. Perhaps this speaks to the mode of the survey influencing the detail in responses (Kiesler and Sproull 1986; Wright 2005). Distasteful language, slang, and stereotyping are all examples of complex ideas or processes that have been condensed. It is possible that survey mode influences this socially undesirable shorthand. In the paper and pencil survey, there was finite space and time to handwrite responses. Students may have endeavored to finish quickly so they could leave class or so that they did not have to write as much. During implementation, online surveys can combat this in two ways. First,

online surveys are taken at the time of the respondent's choosing, making the respondent less likely to rush through the survey. Second, as computers and similar devices are now commonplace devices in many individuals' lives, typing is more comfortable than writing for many individuals. Consequently, condensed, shorter answers may be seen on in-person surveys, whereas web-based surveys may see longer, more detailed responses (Kiesler and Sproull 1986; Wright 2005). This is contrasted with the online environment where respondents could quickly type volumes and be very specific. These possibilities are supported by a simple auxiliary *t*-test. When we counted the number of characters in the open-ended responses, the mean number of characters in the in-person survey was 63 (standard deviation = 45) and the mean for online surveys was 91 (standard deviation = 74). This yielded a *t*-test of 7.10, which is statistically significant. In short, respondents wrote more for the online surveys than on the paper surveys. However, this should be interpreted with caution, as we did not find a difference in response length by undesirable response.

There were limitations to our study that are worth noting. First, our sample lacks generalizability. The prevalence of substantive variations in responses to the stimulus should be minimal in such a constrained sample. Yet, the variability that we saw across responses suggests that this variation, and any differences in survey mode found here, would be replicated in the general population. As such, work of this kind should be replicated using more representative samples. Another limitation of the current study was its reliance on a sample of criminal justice students to interpret neighborhood disorder. A small number of responses included direct references to concepts—such as “broken windows” or “gang affiliated”—from the criminal justice curriculum. However, given that only a few subjects responded in such a way, we were relatively unconcerned about this. Lastly, it may be that our propensity model failed to balance adequately the online and in-person groups and that those who were likely to take online courses (older, white, female) were simply more than or as polite as their in-person student colleagues, though we took careful steps to prevent this imbalance.

Conclusion

This study demonstrates that the mode of administration—whether online or paper and pencil—is important for socially (un)desirable responses, but not in the expected way. We find that online surveys result in reduced rates of socially undesirable responses in the Poisson models, though survey mode is not significant in the Raush models. Additionally, in-person survey respondents were far more likely to offer blatant stereotyping than their online counterparts. In sum, online surveys may not impact response content in open-ended questions. As such, it is plausible that the concern over online surveys inspiring

noxious or malicious responses may be unnecessary. Future research would do well to focus on how perceptions of internet anonymity shape respondents' willingness to offer truthful, sensitive, or socially undesirable responses. Understanding differences in responses and social desirability by mode is vital to planning future surveys, particularly those surveys assessing delicate social topics. For instance, health research is increasingly using web-based surveys, and the field relies heavily on open-ended responses (Hsieh and Shannon 2005; van Gelder, Bretveld, and Roeleveld 2010). With the increased use of qualitative content analysis software (Krippendorff 2012; La Pelle 2004), it is important to understand the effects of mode on response bias.

Finally, future research should be open to exploring how a mixed method perspective facilitates understanding social desirability. This work folds together coding of open-ended questions, visual methods, and quantitative analysis to understand how social undesirability takes shape across survey modes. Without this integration of methods, our study would not be feasible. This is not to say that mono-method studies, specifically quantitative studies, are unable to explore issues of social (un)desirability, but rather, that mixed method studies have the ability to examine respondents reactions to questions that elucidate socially undesirable reactions in ways that multiple choice questions fall short.

References

- Anderson, Elijah. 2012. The iconic ghetto. *The Annals of the American Academy of Political and Social Science* 642 (1):8-24.
- Austin, Peter C. 2011a. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research* 46 (3):399-424.
- Austin, Peter C. 2011b. A tutorial and case study in propensity score analysis: An application to estimating the effect of in-hospital smoking cessation counseling on mortality. *Multivariate Behavioral Research* 46 (1):119-151.
- Belson, William A., and John G. Cleland. 1986. *Validity in survey research: With special reference to the techniques of intensive interviewing and progressive modification for testing and constructing difficult or sensitive measures for use in survey research: A report*. Brookfield, VT: Gower.
- Berrens, Robert P., Alok K. Bohara, 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):1-22.
- Bishop, George F., Alfred J. Tuchfarber, and Robert W. Oldendick. 1986. Opinions on fictitious issues: The pressure to answer survey questions. *Public Opinion Quarterly* 50 (2):240-250.
- Blevins, Kristie R., and Thomas J. Holt. 2009. Examining the virtual subculture of johns. *Journal of Contemporary Ethnography* 38 (5):619-648.
- Branigan, Elizabeth. 2002. But how can you prove it? Issues of rigour in action research. *Stronger Families Learning Exchange Bulletin* 2:12-13.
- Cassidy, Wanda, Margaret Jackson, and Karen N. Brown. 2009. Sticks and stones can break my bones, but how can pixels hurt me? Students' experiences with cyber-bullying. *School Psychology International* 30 (4):383-402.
- Christopherson, Kimberly M. 2007. The positive and negative implications of anonymity in internet social interactions: "On the internet, nobody knows you're a dog." *Computers in Human Behavior* 23 (6):3038-3056.
- Couper, Mick P., and Peter V. Miller. 2008. Web survey methods introduction. *Public Opinion Quarterly* 72 (5):831-835.
- Creswell, John W. 2013. *Research design: Qualitative, quantitative, and mixed methods approaches*. Thousand Oaks, CA: Sage.
- Crowne, Douglas P., and David Marlowe. 1960. A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology* 24 (4):349.

- Davis, Darren W. 1997. The direction of race of interviewer effects among African-Americans: Donning the black mask. *American Journal of Political Science* 41 (1):309-322.
- de Leeuw, Edith D. 2012. Counting and measuring online: the quality of internet surveys. *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique* 114 (1):68-78.
- Edwards, Allen L. 1957. *The social desirability variable in personality assessment and research*. New York: Holt, Rinehart & Winston.
- Fogel, Joshua, and Elham Nehmad. 2009. Internet social network communities: Risk taking, trust, and privacy concerns. *Computers in Human Behavior* 25 (1):153-160.
- Goffman, Erving. 1963. *Stigma: Notes on the management of spoiled identity*. New York: Simon and Schuster.
- Hardaker, Claire. 2010. Trolling in asynchronous computer-mediated communication: From user discussions to academic definitions. *Journal of Politeness Research: Language, Behaviour, Culture* 6 (2):215-242.
- Hargittai, Eszter. 2007. Whose space? Differences among users and non-users of social network sites. *Journal of Computer-Mediated Communication* 13 (1):276-297.
- Herring, Susan, Kirk Job-Sluder, Rebecca Scheckler, and Sasha Barab. 2002. Searching for safety online: Managing "trolling" in a feminist forum. *The Information Society* 18 (5):371-384.
- Hipp, John R. 2010. Resident perceptions of crime and disorder: How much is "bias," and how much is social environment differences? *Criminology* 48:475-508.
- Holbrook, Allyson L., Melanie C. Green, and Jon A. Krosnick. 2003. Telephone versus face-to-face interviewing of national probability samples with long questionnaires: Comparisons of respondent satisficing and social desirability response bias. *Public Opinion Quarterly* 67 (1):79-125.
- Hsieh, Hsiu-Fang, and Sarah E. Shannon. 2005. Three approaches to qualitative content analysis. *Qualitative health research* 15 (9):1277-1288.
- Kiesler, Sara, and Lee S. Sproull. 1986. Response effects in the electronic survey. *Public Opinion Quarterly* 50 (3):402-413.
- Kreuter, Frauke, Stanley Presser, and Roger Tourangeau. 2008. Social desirability bias in CATI, IVR, and web surveys: The effects of mode and question sensitivity. *Public Opinion Quarterly* 72 (5):847-865.
- Krippendorff, Klaus. 2012. *Content analysis: An introduction to its methodology*. Thousand Oaks: Sage.
- Kuran, Timur. 1995. *Private truths, public lies: The social consequences of preference falsification*. Cambridge, MA: Harvard University Press.

- Kuran, Timur, and Edward J. McCaffery. 2004. Expanding discrimination research: Beyond ethnicity and to the web. *Social Science Quarterly* 85 (3):713-730.
- Kuran, Timur, and Edward J. McCaffery. 2008. Sex differences in the acceptability of discrimination. *Political Research Quarterly* 61 (2):228-238.
- La Pelle, Nancy. 2004. Simplifying qualitative data analysis using general purpose software tools. *Field Methods* 16 (1):85-108.
- Latkin, C. A., Daniele German, Wei Hua, and A. D. Curry. 2009. Individual-level influences on perceptions of neighborhood disorder: A multilevel analysis. *Journal of Community Psychology* 37 (1):122-133.
- Mann, Leon. 1981. The baiting crowd in episodes of threatened suicide. *Journal of Personality and Social Psychology* 41 (4):703.
- Maxwell, J. 2005. *Qualitative research design: An interactive approach*. Thousand Oaks, CA: Sage.
- Mayoh, J., and A. J. Onwuegbuzie. 2013. Toward a conceptualization of mixed methods phenomenological research. *Journal of Mixed Methods Research*.
- Miller, Thomas I., Michelle Miller Kobayashi, Erin Caldwell, Sarah Thurston, and Ben Collett. 2002. Citizen surveys on the web: General population surveys of community opinion. *Social Science Computer Review* 20 (2):124-136.
- Phillips, Whitney. 2011. Meet the trolls. *Index on Censorship* 40 (2):68-76.
- Presser, Stanley, and Linda Stinson. 1998. Data collection mode and social desirability bias in self-reported religious attendance. *American Sociological Review* 63 (1):137-145.
- Reise, Steven P., and Tyler M. Moore. 2012. An introduction to item response theory models and their application in the assessment of noncognitive traits.
- Richman, Wendy L., Sara Kiesler, Suzanne Weisband, and Fritz Drasgow. 1999. A meta-analytic study of social desirability distortion in computer-administered questionnaires, traditional questionnaires, and interviews. *Journal of Applied Psychology* 84 (5):754.
- Saldaña, J. 2009. *The coding manual for qualitative researchers*. Thousand Oaks, CA: Sage.
- Sampson, Robert J. 2012. *Great American city: Chicago and the enduring neighborhood effect*. Chicago, IL: University of Chicago Press.
- Sampson, Robert J., and Stephen W. Raudenbush. 2004. Seeing disorder: Neighborhood stigma and the social construction of “broken windows.” *Social Psychology Quarterly* 67 (4):319-342.
- Savage, Charlie, Edward Wyatt, and Peter Baker. 2013. U.S. confirms gathering of web data overseas. *New York Times*, June 7, 2013.

- Schreck, Christopher J., Graham C. Ousey, Bonnie S. Fisher, and Pamela Wilcox. 2012. Examining what makes violent crime victims unique: Extending statistical methods for studying specialization to the analysis of crime victims. *Journal of Quantitative Criminology* 28 (4):651-671.
- Skogan, Wesley G. 1990. *Disorder and decline: Crime and the spiral of decay in American neighborhoods*. Berkeley, CA: University of California Press.
- Smith, Douglas A. 1986. The neighborhood context of police behavior. *Crime and Justice* 8 (Communities and Crime):313-341.
- Stewart, Miriam, Edward Makwarimba, Alison Barnfather, Nicole Letourneau, and Anne Neufeld. 2008. Researching reducing health disparities: Mixed-methods approaches. *Social Science & Medicine* 66 (6):1406-1417.
- Sudman, Seymour, and Norman M. Bradburn. 1974. *Response effects in surveys: A review and synthesis*. Chicago, IL: Aldine Publishing Company.
- Sutton, Robbie M., and Stephen Farrall. 2005. Gender, socially desirable responding and the fear of crime: Are women really more anxious about crime? *British Journal of Criminology* 45 (2):212-224.
- Taub, Richard P., D. Garth Taylor, and Jan D. Dunham. 1984. A theory of reactions to crime and urban neighborhood change. In *Paths of neighborhood change: Race and crime in urban America*. Chicago, IL: University of Chicago Press.
- Taylor, Ralph B. 2001. *Breaking away from broken windows: Baltimore neighborhoods and the nationwide fight against crime, grime, fear and decline*. Boulder, CO: Westview Press.
- Tourangeau, Roger. 2000. *The psychology of survey response*. Cambridge, MA: Cambridge University Press.
- Tourangeau, Roger, and Tom W. Smith. 1996. Asking sensitive questions: The impact of data collection mode, question format, and question context. *Public Opinion Quarterly* 60 (2):275-304.
- Tourangeau, Roger, and Ting Yan. 2007. Sensitive questions in surveys. *Psychological Bulletin* 133 (5):859.
- van Gelder, Marleen M. H. J., Reini W. Bretveld, and Nel Roeleveld. 2010. Web-based questionnaires: The future in epidemiology? *American Journal of Epidemiology* 172 (11):1292-1298.
- Wacquant, Loic J. D. 1993. Urban outcasts: Stigma and division in the black American ghetto and the French urban periphery. *International Journal of Urban and Regional Research* 17:366-383.
- Wallace, Danielle. 2011. A test of routine activities and neighborhood attachment explanations for bias in disorder perceptions. *Crime & Delinquency*.
- Werthman, C., and I. Piliavin. 1967. Gang members and the police. In *The police: Six sociological essays*, edited by D. Bordua. New York: Wiley.

- Wright, Kevin B. 2005. Researching internet-based populations: Advantages and disadvantages of online survey research, online questionnaire authoring software packages, and web survey services. *Journal of Computer-Mediated Communication* 10 (3).
- Young, Alyson L., and Anabel Quan-Haase. 2009. Information revelation and internet privacy concerns on social network sites: a case study of Facebook. Paper presented at the Fourth International Conference on Communities and Technologies.
- Zimbardo, Phillip G. 1969. The human choice: Individuation, reason, and order versus deindividuation, impulse and chaos. In *Nebraska symposium on motivation*. Lincoln, NE: American Psychological Association.

Notes

¹ While we discuss this further in the coming text, it is important to note here that the photograph employed strongly hints at the racial composition of the neighborhood, which in this case is a Hispanic neighborhood. Some discussion of race occurs in over 50 percent of the responses.

² In some cases, students were enrolled in more than one class that was recruited for the study. This number reflects the non-replicated number of students asked to take the survey.

³ The project involved two experiments combined in one instrument. The first, a factorial vignette design, randomly assigned several measures of the code of the street into vignettes, to which subjects were asked to respond. The second experiment involved testing perceptions and interpretations of disorder and how best to measure them. The first experiment in the project was not tested in this study.

⁴ The neighborhood photographed in this study was chosen and photographed with attention. First, we photographed in color with a high megapixel camera rather than black and white since black and white photography adds grittiness to images, which may be confounded with neighborhood disorder. Lastly, the photograph used in the study was from Chicago to reduce the likelihood that respondents would recognize the location.

⁵ A relevant example of this comes from early coding discussions about responses that referred to the community pictured in the stimulus with phrases such as “full of Mexicans” or “run by blacks.” Some coders suggested that statements indicating the community was somehow “overrun” by a minority population were inflammatory and objectionable. Arguing for a more restrictive operationalization of social undesirability, the authors decided that while statements such as these were perhaps unpalatable to some, they may also be ways to express the idea that the area in the picture is populated primarily by an ethnic group. This example shows the conservative nature of coding we employed. As such, phrases such as “full of...” and even “run by Mexicans/Chicanos/Latinos” were not coded as socially undesirable.