



Research in the Sociology of Work

Emerald Book Chapter: A New Approach for Studying Stratification and Religion: Early Results from a National Internet-Based Field Experiment study of U.S. Churches

Michael Wallace, Bradley R.E. Wright, Christine Zozula, Stacy Missari, Christopher M. Donnelly, Annie Scola Wisnesky

Article information:

To cite this document: Michael Wallace, Bradley R.E. Wright, Christine Zozula, Stacy Missari, Christopher M. Donnelly, Annie Scola Wisnesky, (2012), "A New Approach for Studying Stratification and Religion: Early Results from a National Internet-Based Field Experiment study of U.S. Churches", Lisa A. Keister, John Mccarthy, Roger Finke, in (ed.) Religion, Work and Inequality (Research in the Sociology of Work, Volume 23), Emerald Group Publishing Limited, pp. 369 - 397

Permanent link to this document:

[http://dx.doi.org/10.1108/S0277-2833\(2012\)0000023018](http://dx.doi.org/10.1108/S0277-2833(2012)0000023018)

Downloaded on: 21-08-2012

References: This document contains references to 28 other documents

To copy this document: permissions@emeraldinsight.com

This document has been downloaded 2 times since 2012. *

Users who downloaded this Chapter also downloaded: *

(2010), "Research in the Sociology of Work", Christine L. Williams, Kirsten Dellinger, in (ed.) Gender and Sexuality in the Workplace (Research in the Sociology of Work, Volume 20), Emerald Group Publishing Limited, [http://dx.doi.org/10.1108/S0277-2833\(2010\)0000020017](http://dx.doi.org/10.1108/S0277-2833(2010)0000020017)

(2001), "Preface", in (ed.) 8 (Research in Experimental Economics, Volume 8), Emerald Group Publishing Limited, pp. ix - x [http://dx.doi.org/10.1016/S0193-2306\(01\)08001-2](http://dx.doi.org/10.1016/S0193-2306(01)08001-2)

(2010), "Research in the Sociology of Work", Christine L. Williams, Kirsten Dellinger, in (ed.) Gender and Sexuality in the Workplace (Research in the Sociology of Work, Volume 20), Emerald Group Publishing Limited, [http://dx.doi.org/10.1108/S0277-2833\(2010\)0000020018](http://dx.doi.org/10.1108/S0277-2833(2010)0000020018)

Access to this document was granted through an Emerald subscription provided by ASA Delegate

For Authors:

If you would like to write for this, or any other Emerald publication, then please use our Emerald for Authors service. Information about how to choose which publication to write for and submission guidelines are available for all. Please visit www.emeraldinsight.com/authors for more information.

About Emerald www.emeraldinsight.com

With over forty years' experience, Emerald Group Publishing is a leading independent publisher of global research with impact in business, society, public policy and education. In total, Emerald publishes over 275 journals and more than 130 book series, as well as an extensive range of online products and services. Emerald is both COUNTER 3 and TRANSFER compliant. The organization is a partner of the Committee on Publication Ethics (COPE) and also works with Portico and the LOCKSS initiative for digital archive preservation.

*Related content and download information correct at time of download.

A NEW APPROACH FOR STUDYING STRATIFICATION AND RELIGION: EARLY RESULTS FROM A NATIONAL INTERNET-BASED FIELD EXPERIMENT STUDY OF U.S. CHURCHES

Michael Wallace, Bradley R. E. Wright,
Christine Zozula, Stacy Missari, Christopher
M. Donnelly and Annie Scola Wisnesky

ABSTRACT

Purpose – In this chapter, we introduce the Internet-based field experiment (IBFE) that offers numerous advantages for bringing stratification processes “back into” the study of religion. We present preliminary results from a study of class and race discrimination using this approach.

Design/Methodology/Approach – Using names of fictitious characters, we sent e-mails to a nationally representative sample of 4,680 U.S.

Religion, Work and Inequality

Research in the Sociology of Work, Volume 23, 369–397

Copyright © 2012 by Emerald Group Publishing Limited

All rights of reproduction in any form reserved

ISSN: 0277-2833/doi:10.1108/S0277-2833(2012)0000023018

Christian churches asking about possible membership. The e-mails varied only in the perceived race and class of the senders. We utilize a mixed methods approach to analyze variation in the content of the church responses.

Findings – Our early findings suggest significant variation by race/class manipulation, religious denomination, and region of the country in churches' responses as well as the length of time they took to reply, the length of the response, the warmth, religious tone, and several other dimensions.

Research limitations/Implications – This study raises new opportunities for Internet-based research on religion in a variety of social settings, but there is not yet a well-established set of “do’s” and “don’ts” for how to proceed. We advocate the development of a protocol of best practices as this research method develops.

Originality/Value – This study demonstrates the opportunities and pitfalls of the IBFE and the advantages it provides for studies of stratification and religion. Ours is the first study to apply this emerging method to the study of religion and stratification.

Keywords: Internet-based field experiment; audit study; discrimination; religious denominations; class; race

INTRODUCTION

In this chapter, we present a new methodological approach for incorporating dimensions of stratification into the study of religion and present preliminary results from a study based on a national sample of 4,680 U.S. churches. Specifically, we use the *Internet-based field experiment* method to determine levels of class and race discrimination displayed by church representatives toward potential new members, and we analyze this discrimination as it varies nationally, by religious denomination, by characteristics of the churches, and by characteristics of communities in which churches are located. Field experiments have been widely used to study discrimination by race, class, gender, and sexual orientation, but few field experiments have focused on topics of interest to scholars of religion. In recent years, researchers have begun to use the Internet to conduct field experiments, but this trend has not yet caught on among scholars studying

religion. As we demonstrate, Internet-based field experiments are amenable to the study of religion topics; indeed, they can help disentangle complex issues of causality revealed in previous research on religion, and they raise new questions that have been previously overlooked. Although the research we describe here focuses specifically on class and race discrimination by churches, the method is adaptable to a broad range of stratification dimensions and institutional settings in which religion might be studied.

This chapter has three objectives. First, we provide a detailed discussion of the Internet-based field experiment (hereafter, IBFE) method used in our study. As we will show, the IBFE method lends itself to collecting a rich mixture of quantitative and qualitative data (Paluck, 2010). This discussion provides religion scholars with an understanding about the method as well as ideas as to how it can be used in future research. Second, we give an overview of the early results emerging from our study. We will also discuss some lines of inquiry that we anticipate pursuing with these data in future research. Third, we assess the opportunities and pitfalls of the IBFE method for future stratification-based studies of religion. We also discuss potential extensions and alternative applications of this method.

AN OVERVIEW OF THE INTERNET-BASED FIELD EXPERIMENT METHOD

Festinger and Katz (1953, p. 137) succinctly described the field experiment as “a theoretically oriented research project in which the experimenter manipulates an independent variable in some real social setting in order to test some hypothesis.” Field experiments incorporate many features of conventional laboratory experiments but are conducted in natural settings. Like laboratory experiments, field experiments use random assignment of subjects into control and experimental groups and/or matching to study the effect of a single intervention on a particular outcome variable. The experimental design typically assures that potential alternative variables are controlled, so that the researcher can isolate the effects of the experimental manipulation on the outcome variable. Because field experiments are conducted in natural settings, they avoid the artificiality of conventional laboratory experiments and achieve a higher level of external validity. Thus, the results of field experiments are more easily extrapolated to real life situations in society. Also, unlike conventional experiments, subjects in field experiments are typically unaware that they are involved in a research study.

Consequently, field experiments are less prone to contamination by demand characteristics of the experimental situation in which subjects' behaviors are altered by attempts to either please or frustrate the experimenter. As disadvantages, field experiments are usually seen as providing less control over extraneous variables in the environment and less replicability than conventional experiments, but these disadvantages can be partially mitigated with the IBFE method.

A large and growing genre of field experiments conducted primarily by sociologists and economists is the audit study that is designed to measure discriminatory behavior based on race, sex, or some other status characteristic. Audit studies date back to LaPiere's (1934) classic study of discriminatory attitudes versus discriminatory behavior. LaPiere toured the country for two years with a Chinese couple. They visited 251 hotels and restaurants and were refused service only once. Later, LaPiere sent questionnaires to the same restaurants and hotels and found that 92% of establishment owners said they would refuse service to "members of the Chinese race." In short, these proprietors were more discriminatory in their attitudes than in their actual behavior. Some 70 years later, [Pager and Quillian \(2005\)](#) devised a similar study to examine employers' attitudes about hiring ex-offenders and their actual behavior. In surveys, they found that employers typically professed a willingness to hire ex-offenders, but in practice the callback rate for ex-offenders who actually applied was low. Their findings reverse LaPiere's findings to indicate that employers are more discriminatory in behavior than attitudes.

In recent years, audit studies have focused on discrimination in the housing market ([Carpusor & Loges, 2006](#); [Massey & Lundy, 2001](#)), the labor market ([Bertrand & Mullainathan, 2004](#); [Pager, Western, & Bonikowski, 2009](#); [Riach & Rich, 2006](#)), and consumer transactions ([Dovidio, Gluszek, John, Ditlmann, & Lagunes, 2010](#)). Most of these audit studies investigate discrimination by race, ethnicity, or sex, but a few have looked at other dimensions. For instance, [Ahmed, Andersson, and Hammarstedt \(2008\)](#) studied discrimination against lesbians in the housing market. [Pager \(2003\)](#) studied the impact of having a criminal record on hiring in entry-level positions (see also [Pager & Quillian, 2005](#)). [Rooth \(2009\)](#) found that obesity and attractiveness led to differential callback rates for job applicants in the Swedish labor market. Finally, [Correll, Benard, and Paik \(2007\)](#) combined an audit study with a traditional laboratory experiment to examine the "motherhood penalty" for women who applied for jobs. The laboratory component showed that women who are mothers are viewed as less competent and less hireable and would be offered a lower starting salary than

women without children. The audit component demonstrated that mothers received about half as many callbacks as women without children whereas men experienced no penalty for being parents.

In contrast to audit studies that involve person-to-person interaction, an increasing share of field experiments take the form of correspondence studies in which some form of communication such as resumes, phone calls, or e-mails, is used. Our current study of religious discrimination qualifies as a correspondence study. Correspondence studies overcome some of the limitations of audit studies. First, they do not require the extensive training of auditors to reduce differences in interpersonal presentation, particularly in response to perceived discrimination (Heckman, 1998). Second, correspondence studies overcome geographic limitations of audit studies that provide potentially broader representativeness of samples. Finally, in labor market studies, in particular, correspondence studies permit access to a wider range of jobs than are typical of traditional audit studies. Extrapolating to our own study, with the help of the Internet, we contacted 4,680 separate churches nationwide.

In the Internet age, it seems natural to develop field experiment methods, including but not limited to audit/correspondence studies, that take advantage of Internet-based resources. The Internet opens up limitless opportunities for conducting field experiments on inequality in a variety of institutional settings, yet Internet-based studies are rare.¹ Pager's (2007) comprehensive review of field experiments analyzing employment discrimination failed to identify a single study that used the Internet. In a special journal issue on field experiments in economics, Lahey and Beasley (2009) describe the advantages of "computerizing audit studies" of employment discrimination, but they cite no published research using the Internet. Other recent reviews of field experiments also fail to mention any publications featuring Internet-based designs (Harrison & List, 2004; Paluck, 2010). Nonetheless, Ahmed and his colleagues have generated several Internet-based audit studies of housing discrimination in Sweden (e.g., Ahmed et al., 2008; Ahmed & Hammarstedt, 2009).

We could find no previous IBFE in the sociology of religion, and we could only find a single audit study that investigated religious-based discrimination (Banerjee, Bertrand, Datta, & Mullainathan, 2009). An exhaustive search of the literature revealed only a handful of conventional field experiments related in any way to the sociology of religion. For instance, Johnson (1971) conducted a field experiment to examine religious change among subjects attending a Billy Graham crusade. Carey (1971) investigated the role of peer influence in daily attendance at mass among schoolchildren.

Soetevent (2005) conducted a field experiment in 30 Dutch churches to show that “open” offerings elicited 10% higher contributions than “closed” offerings. Malhotra (2010) found that religious persons are more likely than nonreligious persons to respond to appeals for charity on days when they attend religious services, but no more likely to do so on other days. Finally, in the only religion-based audit study we could find, Banerjee et al. (2009) examined the role of caste and religion in employment discrimination in India. They sent out 3,160 fictitious resumes to 371 job openings in the IT sector in New Delhi. They found there was some discrimination by caste for call center jobs, but not software jobs, and there was no discrimination against Muslims for either type of job.

This paucity of religion-based field experiments, particularly those that deal with dimensions of stratification, is perhaps not surprising given recent calls to bring stratification processes “back into” the study of religion (Edwards, 2008; McCloud, 2007; Smith & Faris, 2005). Nevertheless, IBFEs, particularly audit/correspondence studies, have much potential to invigorate stratification-based studies of religion. We describe our study as an example of how this can happen.

METHOD AND RESEARCH DESIGN OF THE PRESENT STUDY

Our study revolves around a very simple research question: What happens when potential members inquire about joining a new church? Are they treated differently by church officials based on their perceived race or class? Posing as an adult male with family who is contemplating a move to a new area of the country and interested in joining a new church, we contacted churches to solicit information about joining. Our contact with the churches was initiated with the following e-mail (*adapted for Catholics with the italicized parentheticals*):

Dear Pastor (*Father*),

My family and I are planning to move to your area in the next several months. We're looking for a new church (*parish*) to attend, and we're hoping to get some information about your church (*parish*). We'd like to find a church (*parish*) similar to the one that we regularly attend now. So would you please let me know how big your congregation (*parish*) is, your service times, and any other additional information that may be helpful?

Thank you very much,

(Signature)

The subject line of the e-mail was “Moving to your area.” We identified the degree of discrimination based on quantitative and qualitative patterns in the responses received from church officials.

The key manipulation in the study is the perceived race or class of the e-mail authors, whom we labeled “characters.” We designated the race of each character on the signature line by utilizing given names and surnames that signaled likely racial or ethnic identities of the characters as follows: “white” (Scott Taylor, Greg Murphy); “black” (Jamal Washington, Tyrone Jefferson); “Hispanic” (Carlos Garcia, José Hernandez); or “Asian” (Wen-Lang Li, Jong Soo Kim). We created two names for each racial/ethnic identity to offset idiosyncratic responses to certain names. Among white characters, in addition to the “regular” whites, we also created perceived class distinctions by adding the following phrases at the end of the first sentence of the e-mails (shown in italics): “rich whites” (“My family and I are planning to move to your area in the next several months *as I’ve recently sold my company*”) and “poor whites” (“My family and I are planning to move to your area in the next several months *as I’ve recently lost my job.*”). In this way, we created a standard set of 6 race/class manipulations among our 12 characters as shown in Fig. 1. This configuration builds in two dimensions of comparison: by race (white control/black/Hispanic/Asian) and by class (white control/white rich/white poor) with the white control identity serving as the reference group in both comparisons. While we plan to analyze these two dimensions separately in future papers, here we combine them into one experimental condition that we call the *race/class manipulation*.

White (Control)	White (Rich)	White (Poor)
Black		
Hispanic		
Asian		

Fig. 1. Race and Class Manipulation of Senders in the Sample.

We created separate name-identified e-mail accounts for each of the 12 characters. With the assistance of several graduate and undergraduate assistants, we sent the e-mails described above to a nationally representative sample of churches between May and July 2010. These e-mail accounts were then used to collect the responses of church officials. We received many responses on the same day we sent the e-mails; in other cases, responses came several weeks later, often with an apology for the slow response. The length of time it took to receive a response is one of the key outcome variables.

We are interested in whether the pattern of responses by church officials varies according to the perceived race or class identity of the characters. We developed three primary types of data to measure the pattern of response. First, we measured whether the characters received a response at all from the pastor, church secretary, or some other representative of the church. We also recorded if characters were contacted more than once by a single church as this might indicate more sincere interest. Since many churches seek to increase member contributions, we suspect that rich white characters might receive more second contacts than others. By aggregating responses across the sample, we create quantitative response rates by race and class. Second, we assessed the quality of the response as judged by its length, its warmth, its religious tone, the quality of information provided, the length of time to reply, and several other dimensions. We conducted a content analysis of the churches' responses to the e-mails by creating a survey with approximately 70 questions.² With the help of graduate assistants, we used the survey instrument to "interrogate" the content of the e-mails and coded this content. We held several training sessions to familiarize our assistants with the coding and had frequent meetings to assure consistent procedures were followed. Finally, we assembled a qualitative data set comprised of representative responses from church representatives. For the content analysis and qualitative data collection, the e-mails were scrambled and characters' names were redacted so that coders could not ascertain the race or class of the original characters. We posit that systematic variation by race or class in any of these three types of indicators indicates discriminatory treatment by church representatives.

The Internet-based nature of our field experiment means that there were no geographical barriers to the study. Thus, in order to assure a nationally representative sample, we employed a sampling frame representative of different geographic regions in the United States as well as a range of major Christian denominations. To achieve geographical representation, we obtained a list of Congressional districts and arrayed them in inverse order

according to the population density of the districts. We then used interval sampling to select 65 districts (approximately 1 out of 7) for inclusion in the sample. In addition to being geographically dispersed, these districts encompassed a range of settings including rural, urban central city, suburban, and small town locations.

We then identified 12 major, organized Christian denominations in the United States including Catholic; Mainline Protestant (Methodist, Episcopalian, Evangelical Lutheran, Presbyterian, and American Baptist); and Evangelical Protestant (Southern Baptist, Assembly of God, Lutheran Missouri Synod, Church of Christ, Willow Creek Association, and Pentecostal churches) traditions. For most denominations, we were able to locate a website that provided a comprehensive national list of churches. For denominations with no readily accessible websites, we used the Google search engine to identify churches. By either method, we ultimately identified churches by name along with their locations, contact information, and (in many cases) church websites. One key requirement for inclusion in the sample was that the church must have a working e-mail address by which our characters could contact them; these e-mail addresses might be for the church pastor or priest, a church secretary, or an anonymous church address. Using this information, we attempted to locate 6 churches (one for each race/class manipulation) for each of the 12 denominations within each Congressional district. We attempted to identify these 72 churches within a reasonable radius of a central location (such as a major city or town) within each district. With these parameters, we sometimes had to extend the radius some distance from the central location and in some cases our search spilled over into adjacent Congressional districts. This led to a few duplicate churches being included in the original sample (the same church was chosen from two different districts); these duplicates were weeded out and replaced with nonduplicates. In addition, some of our initial e-mails bounced back to us or were answered by someone no longer affiliated with the target church. In these cases, we either attempted to contact a second e-mail affiliated with the same church or chose a different church altogether.

This design yielded a nationally representative sample of 4,680 churches (65 Congressional districts \times 12 denominations per district \times 6 race/class manipulations per denomination). In each district, the 6 race/class manipulations were assigned randomly to the 6 churches within each denomination. In addition, the specific names of each character within each race/class manipulation (e.g., José Hernandez or Carlos Garcia) were assigned by matching to exactly half of the cases within each denomination. By randomizing each aspect of the design, we effectively control for most

plausible alternative explanations of differential treatment. In order to assess race and sociodemographic data about the church, we examined church websites when available. Website data were available for about 85% of the churches in our sample. Undergraduate assistants were instructed to view written and visual content of the websites and to code the following variables: gender and perceived race of the pastor or priest, perceived racial/ethnic composition of the congregation, number of staff, and quality of website. Race composition of church was measured using five broad categories: predominantly white, black, Hispanic, Asian, or racially diverse. Number of staff was determined by counting the number of persons named on the “church staff” page and is used as a proxy for bureaucratization of the church. Quality of website was measured using three categories (below average, average, above average) and was used as a proxy for church affluence. Our assistants were trained to code these variables accurately and consistently and they reported relatively few problems in the coding process. Among other things, these data allowed us to compare the characteristics of churches that responded to those that did not. They also enabled us to ascertain how race and sociodemographic characteristics of churches intersected with the race and class of our characters to determine the pattern of response.

In order to address the race and sociodemographic composition of the communities in which churches are located, we geocoded data from the 2010 Census to our church data set matched to the zip codes of the churches. Fortunately, the data collection phase of our project took place while the Census was being conducted. The Census data included measures of variables such as race composition, median income, poverty rates, and average education levels of the communities. These data were also collected for all churches in the sample, whether or not the churches responded to the e-mails. While zip code level data might mask some variability at lower levels such as neighborhoods, they still constitute relatively homogenous geographic units and represent a reasonable catchment area for local church congregations. These two sources of socio-geographic data represent an added dimension of the IBFE method. They also address one of the major concerns about conventional field experiments – that such designs are unable to control for extraneous variables in the field that might affect the outcome variable.

In summary, our design yielded a multidimensional data set that incorporates both quantitative and qualitative data and lends itself to a multimethod approach. Thus, there are multiple avenues by which to pursue our central research question about whether and under what conditions church officials discriminate against potential congregants based on race or class.

SOME EARLY RESULTS AND FUTURE LINES OF INQUIRY

In this section, we provide some early results from our study. First, we discuss results from a quantitative analysis of response rates in which we examine the effects of the race/class manipulations and denomination. Second, we show the results for a quantitative analysis of data from the content analysis portion of the study that examines the effects of the race/class manipulation on selected characteristics measuring the quality of responses. Third, we provide some illustrative data from the qualitative portion of the study, in particular, a preliminary analysis of gatekeeping. Finally, we provide a brief discussion of some future lines of inquiry for this study.

Quantitative Analysis of Response Rates

This section's results speak to two questions: (1) To what extent do the response rates of churches vary by the perceived race or class of the writer? (2) To what extent do response rates differ by denomination of the church? These analyses address only *whether* we received a response and do not consider the *quality* of the response. Also, we address only first responses here, and do not include analyses of multiple responses that occurred in some cases.³ One attractive feature of our IBFE design is that extraneous factors that might affect response rates are controlled by the random assignment of churches to the different experimental conditions. Thus, we can efficiently isolate the effects of race/class and church denomination with a two-way analysis of variance (ANOVA).

While we collected data from churches in 12 different Christian denominations, for brevity's sake we analyze here only the 5 Mainline Protestant denominations: Episcopalian, Evangelical Lutheran, American Baptist, Methodist, and Presbyterian. We sent e-mails to 390 churches in each of these denominations, resulting in a sample size of 65 (390/6) churches per experimental condition in each denomination. On average, 60% of the Mainline churches responded at least once to the e-mails. Episcopalian churches were the most industrious respondents, answering 69% of the e-mails they received. Evangelical Lutherans followed at 62%, and the remaining three responded at rates between 55% and 59%.

As shown in [Table 1](#), there is substantial variation in response rates by race and class. Among all Mainline churches, 67% replied to the white

Table 1. Two-Way Analysis of Variance Model Predicting Response Rates from Churches by Race/Class Manipulation and Denomination.

Denomination	Race/Class Manipulation						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	White (Control)	White (Rich)	White (Poor)	Black	Hispanic	Asian	All
Episcopalian	.82	.63	.74	.71	.65	.63	.69
Evangelical Lutheran	.75	.68	.75	.52	.57	.42	.62
American Baptist	.65	.55	.54	.72	.57	.49	.59
Methodist	.66	.58	.60	.46	.57	.51	.56
Presbyterian	.48	.71	.62	.57	.52	.40	.55
All Mainline Churches	.67	.63	.65	.60	.58	.49	.60

Note: Model is significant at $p < .000$; sample size is 64–65 churches per cell; cells report the percentage of churches who responded.

control group. About the same percentage, 63% and 65%, responded to the two other white conditions, the white rich and white poor respondents. In contrast, fewer responded to the black and Hispanic respondents, at 60% and 58%, respectively. The lowest response rate, 49%, was to the e-mails with Asian names. Accordingly, the simple response rate for Asians was 73% (.49/.67) that of the white control group. The difference in response rates across the two experimental conditions was statistically significant with an $F = 5.78$ ($p = .000$).

Another way to examine these data is to look at whether the five Mainline denominations differ from each other in how they responded by race and class. Table 1 also shows the response rates for each denomination, and there are meaningful differences. The Episcopalian churches responded most frequently to e-mails from the white control group, answering 82%. In contrast, they responded to the remaining conditions at relatively low levels, ranging from 63% to 74%. Evangelical Lutherans showed a strong preference for the e-mails from whites, regardless of class, responding to 68% to 75% of them, but they responded to far fewer black (52%), Hispanic (57%), and, especially, Asian (42%) e-mails. American Baptist churches showed a more mixed pattern. They actually responded most frequently to the black e-mails, at 72%, followed by the white control group, at 65%, and the remaining four groups ranged between 49% and 57%. Methodist churches responded most frequently to the white control group, and least frequently to

the black e-mails. Finally, Presbyterians responded relatively infrequently to e-mails from the white control group, but they responded most frequently to white rich and white poor e-mails. They had the lowest response rate among all Mainline Protestants with Asians, at a meager 40%. A two-way ANOVA model found that the interaction effect between denomination and the experimental condition was statistically significant at $F = 1.84$ ($p = .013$).

In analyses not shown, we also estimated a two-way ANOVA between the experimental condition and the region of the country in which the churches were located. We placed the churches into one of nine regions: New England, Middle Atlantic, East North Central, West North Central, South Atlantic, East South Central, West South Central, Mountain, and Pacific.⁴ In this model, the effect of the experimental condition was statistically significant, following the pattern described above, and the effect of region approached conventional significance levels at $F = 1.78$ ($p = .077$). The churches in the South responded to e-mails least frequently, including the South Atlantic (55% response rate) and East South Central (51%). The regions with the highest response rates varied between 60% and 66%. Also we found a significant interaction between the race/class manipulation and region in church responses, $F = 1.84$ ($p = .013$).

Content Analysis of the Quality of Responses

The foregoing analysis centers on a dichotomous dependent variable measuring whether or not churches responded to our e-mailed inquiries. In this regard, it is similar to previous field experiments that examine outcomes such as responding to a job advertisement, granting a mortgage application, or stopping for a person hailing a taxicab (Pager, 2007). However, the IBFE method can offer the added advantage of providing textual material that is subject to analysis. These data provide rich diversity in how church officials responded, how they talked about their church, and how they interacted with the characters who wrote the e-mails. With the help of four valiant graduate assistants, we content analyzed over 2,700 e-mails that we received from church officials. These assistants used a survey of approximately 70 questions to “interrogate” the content of these e-mails and code the responses. The e-mails were scrambled and the character names were redacted so that the assistants were not aware of the race/class category of the e-mails they coded. To our knowledge, ours is the first correspondence study to couple content analysis with simple quantitative measures of response rates.

In *Table 2*, we show the results of an analysis of several selected characteristics from the content analysis portion of our study. We include responses from all 12 denominations and conduct 7 one-way analyses of variance to assess the impact of the race/class manipulation on 7 characteristics measuring the quality of response. Our null hypothesis was that there are no differences across experimental conditions – that church messages would use similar words, adopt a similar tone, and be of similar length regardless of whom they were responding to.

What we found, however, were systematic differences across the experimental conditions, and these differences mirrored the response-rate analysis in *Table 1* in representing an overall less welcoming attitude to racial minorities and, in some cases, the poor.

Column 1 in *Table 2* presents analysis of the initial greeting in the e-mails sent by the churches. We coded the initial greeting into four categories: informal (e.g., “Hey Carlos”), formal (e.g., “Dear Mr. Garcia”), religious (e.g., “Brother Garcia”), or none. *Table 2* reports the percentage of respondents who received informal greetings, the most common type of greeting and, by our reading, the most welcoming. We interpret the informal greetings as indicating a higher comfort level and friendliness. As shown, whites were significantly more likely to receive informal greetings, with 51% receiving a “hey,” “how are you doing,” or something similar. On the other hand, Asians (36%) and Hispanics (39%) were much less likely to receive such a greeting. Blacks, poor whites, and rich whites fell in between these two extremes.

Table 2. Selected Characteristics of Churches’ Responses to Characters’ E-mails by Race/Class Manipulation, One-Way Analysis of Variance.

Race/Class manipulation	Dependent Variables						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Informal greeting	Religious closing	Word count	How we worship	Warmth of message	Selling the church	Quality of information
White (Control)	.51	.54	172	.42	.08	.07	.09
White (Rich)	.46	.52	176	.39	.00	.06	.06
White (Poor)	.46	.56	161	.41	-.04	-.09	-.06
Black	.42	.50	163	.37	.05	.06	.05
Hispanic	.36	.48	150	.31	.01	-.06	-.08
Asian	.39	.45	143	.31	-.11	-.05	-.08
All	.43	.51	162	.37	.00	.00	-.00
Significance level	.000	.010	.000	.000	NS	.027	.021

Column 2 presents the percentage of church e-mails that gave a religious closing. As with the greeting, we coded the closing into four categories: informal (e.g., “See you around”), formal (e.g., “Sincerely”), religious (e.g., “May God bless you”), or no closing statement. The religious closing was the most common closing, and we felt it indicated closer spiritual affinity between the respondent and the character. Here, we find that the white characters, regardless of their class, received 54% or higher religious closings. In contrast, the racial minorities were all at or below 50%, with Asians receiving 45% religious closings. The magnitude of difference here is not as great as with the initial greetings, but it is in the same direction, with churches putting their best foot forward for whites. Curiously, the modal pattern was to give an informal greeting and a religious closing; whites were most frequently the recipients of this response.

Column 3 presents the total word count of the message sent including the greeting, body of the message, and closing, but not including electronic signatures or attachments. This variable is perhaps the most direct measure of church efforts to build a relationship with the character and the least likely to reflect coding bias. Generally, wordier e-mails indicate more time and effort put into the e-mail by the church official, and this could be interpreted as a proxy for their openness to having the recipient join their church. As shown, e-mail responses to the white control group averaged 172 words. By comparison, e-mails to rich whites averaged 176 words; blacks, 163 words; poor whites, 161 words; Hispanics, 150 words; and Asians averaged only 143 words – about 81% that of rich whites.

Column 4 shows data on “How we worship,” a variable that reflects whether the respondent volunteered information about the style of worship, the type of music, or other religious practices that typically occur in the church. Typically, this type of disclosure represented efforts by the church to seem more open and welcoming to the character (although it sometimes indicated an aspect of gatekeeping – see below). Here, we find that the white control group received the highest percentage at 42%, followed closely by white poor at 41% and white rich at 39%. Blacks received this information about 37% of the time and Hispanics and Asians lagged far behind at 31% each. Thus, Hispanics and Asians received only 74% of the level on this variable as the white control group.

Columns 5, 6, and 7 in [Table 2](#) present three summary ratings made by our coders. After reading each e-mail, the coders were instructed to rate it on several dimensions. Here we show three of those dimensions: the warmth of the e-mail’s tone, whether the letter writer appeared to be explicitly “selling” the church, and the overall quality of information conveyed to the character.

While the coders used five-point scales, the distributions of their ratings varied, so we expressed them as standardized z -scores based on coder-specific means and standard deviations. As shown in column 5, the white control group received the warmest e-mails, and Asians received the least warm, but this was the only outcome in which the race/class differences were not statistically significant. On the other hand, the race/class differences for the selling the church measure in column 6 are statistically significant. The white control group and rich whites received e-mails that most actively sold the church to them. In contrast, poor whites rated extremely low, even below all three racial minorities. Finally, in column 7, we see that the results for quality of information closely parallel those for selling the church. The white poor group received poorer quality information, comparable to Asians and Hispanics. Overall, we note that six of the seven characteristics in [Table 2](#) showed significant race/class variation in responses and the pattern is highly consistent across these variables.

Analysis of Qualitative Data

The content analysis of textual material represents a major advance over previous correspondence studies. However, there are limitations to this analysis. For instance, while we have demonstrated that there are meaningful differences by race/class in the length of the e-mails, these differences may obscure the warmth conveyed by a short, heartfelt message. In addition, qualitative analysis is required to evaluate the subtlety, nuance, and context of messages. For example, in future analyses, we will examine the length of time that it took for church officials to respond to the characters' e-mails. But qualitative analysis of the e-mails reveals that "late" responses were frequently prefaced with an apology for being late followed by a warm, effusive message. Thus, the IBFE nature of this study lends itself to a high-quality, mixed methods analysis where qualitative and quantitative data are mutually reinforcing ([Paluck, 2010](#)).

These qualitative data will likely open new avenues of investigation into patterns by treatment effect, denomination, region of the country, and so forth. For example, the type of information provided in the e-mails gives us the opportunity to examine how churches present themselves to newcomers. This can give us insight into what churches look for as well as what they can offer potential members. These e-mails often contained discussions about what they wanted and expected in members of the church, such as adherence to specific religious doctrine and service to community members. In terms of

what the church could offer new members, some e-mails emphasized their strong faith and style of worship, others talked about the wonderful people, and some discussed programs such as Bible studies, picnics, youth groups, and book clubs.

One fruitful avenue for qualitative analysis is the practice of gatekeeping, which we observed in several e-mails. Gatekeeping occurs when the respondent attempts to screen the character to make sure that the character fits into the church community. One of the clearest forms of gatekeeping occurred when the respondent interrogated the character about his home church, hometown, or himself, in a curt and unfriendly manner. In many cases, this served as a screening process in which the respondent withheld information until the character answered some questions about himself. The respondent effectively redefined the terms of the interaction in order to gain the upper hand. The following examples illustrate this form of gatekeeping:

Hello,
Thank you for your email; however, can you tell us more about you and your family, and where you are coming from? Thank you.
(First Initial, Middle Name, Last Name)
[Scott Taylor, white poor]

Dear Jong Soo Kim,
It would be easier if you told me what type of parish you are coming from and what you are looking for. Are you familiar with the Episcopal/Anglican Church at this time?
(First Name)
[Jong Soo Kim, Asian]

Where are you relocating from?
[Scott Taylor, white poor]

Other common expressions of gatekeeping were what we labeled “brush-offs” – brusque referrals to the church website with little additional information, such as these:

God Bless You!
The best way to get the details regarding the church is to visit our website, [http://www.\(Church Name\).org](http://www.(Church Name).org)
[José Hernandez, Hispanic]

Dear Mr. Washington,
You are welcome to check the information you need from our website: [http://www.\(Church Name\).org](http://www.(Church Name).org).
Thank you for your attention.

In Him,
 (First Name)
 [Jamal Washington, black]

I'm (*First Name, Last Name*) the church secretary we have two service 8 am and
 10:30 am sunday school 9:00 am look at our website
[http://www.\(Church Name\).org](http://www.(Church Name).org)
 [Tyrone Jefferson, black]

These brush-offs convey minimal effort in establishing rapport with the character – the last e-mail fails to even employ minimal punctuation and capitalization – and put the burden of continued interaction back on the character himself. These e-mails are perfunctory to the point of being rude and, if this small sampling is indicative, are disproportionately directed toward nonwhites. Thus, while these e-mails are not overtly discriminatory, they establish a sort of boundary maintenance that signals an unwelcoming environment. Short of not responding at all, e-mails such as these evidence high levels of covert discrimination. In future analyses, we would like to adjust the overall response rates by subtracting out these strongly negative instances of gatekeeping. We anticipate that this will create even sharper differences in the race/class variations in response rates.

Another form of gatekeeping is “theological gatekeeping.” These e-mails tend to be longer and give a detailed description of the church’s beliefs and practices. We believe many of these e-mails attempt to screen characters by making sure that the character’s theology aligns with the church’s. For brevity’s sake, we excerpt portions of longer e-mails that illustrate the concept of theological gatekeeping:

... We are a spiritual organization, which is all a church was ever supposed to be. We do not offer programs of social reform; our benevolent work is limited; and we offer no recreation and entertainment at all. The gospel is our only message; the truth our only commodity. We exist to serve a world of people who are lost and dying in sin. We have nothing to sell; everything we offer is FREE to all. We do not desire to have those we might be able to help to feel any obligation to us. We only want to serve
 [Scott Taylor, white control]

Dear Mr. Murphy
 ... I hope I have not scared you away but we know that our church is not for everyone. We do not have modern music, we do not have large modern facilities, and we are not involved in social justice issues. We do provide a place in the community for many who have not been accepted in other churches.
 Shalom,
 Pastor (First Name)
 [Greg Murphy, white poor]

Other respondents demonstrate practical gatekeeping, as with language. For example:

Greg,
Thank you so much for being interested in our church. Not sure if you know, we are a Cantonese speaking church and it might not be fitting for you and your family.
Thanks!!!
(First Name Last Name)
Church Council Moderator
[Greg Murphy, white control]

These e-mails demonstrate boundary-keeping along racial lines. The respondent does not include information about the language of the church along with other details about the church. The respondent's assumption about the racial identity of the character is clearest in these examples. It is interesting as the respondent's racialization of the character based on name does not appear to be considered problematic as it is depoliticized as a concern over language rather than a concern of racial gatekeeping.

Finally, it is important to note that despite sending out almost 5,000 e-mails over the course of the study, we found little evidence that the churches suspected they were part of a research study. The only e-mail that seemed to raise suspicion about the motives of our inquiries was this one which reads in part:

Dear Greg:
... If you really haven't lost your job and this is just a spam letter to get information, how do you expect God to bless you in a lie? If this is a lie, then let your testimony of losing your job come to pass If this email is the truth, let me know what jobs you are looking for and I'll help you as best I can.
Sincerely
(First Name)
[Greg Murphy, white poor]

Future Lines of Inquiry

In the foregoing sections, we have provided a sampling of analyses we have conducted so far. In this section, we discuss several veins of research we hope to mine in the future with these data.

First, we identify several questions we would like to pursue with the qualitative data. For instance, we would like to more rigorously examine the complicated issue of gatekeeping and whether gatekeeping is more likely to emanate from pastors or church staff, men or women responders, large

churches or small churches, or in response to same-race or different-race characters. We would also like to explore in greater depth how church theology (i.e., “how we worship”) is discussed by church respondents and used to “sell the church.” In a related sense, we would like to investigate expressions of diversity (or lack thereof) regarding the church congregation. And we contemplate a project exploring the gendered nature of responses as reflected in male and female church officials.

Second, we look forward to analyzing the website data in greater detail. With these data, we can answer several questions, including: How are church responses to characters of different races affected by the race of the pastor or the race composition of the congregation? How do churches that responded to characters’ e-mails differ from churches that did not respond and how does the pattern of responses differ by the race of the characters? Does the affluence of the church affect the pattern of responses to characters of different classes? Does the bureaucratic structure of the church affect the response pattern? Finally, does the pattern of response differ for churches that do not have websites?

Third, the geocoding of zip code data permits us to address broader questions about the socio-geographic context in which churches are located. With these data, we can more adequately address the influence of contextual variables such as social class composition, education levels, and poverty rates in the community and how that affects response rates for characters of different classes. We can also investigate how these factors along with race composition of the community reinforce or mitigate the effects of race composition of the churches on response patterns. With the Census question about change in residence, we can investigate the impact of neighborhood residential turnover on response patterns.

All of these questions can be pursued with either the full sample or within subsamples. For instance, we have the opportunity to explore more fully whether response patterns differ according to region of the country. We are excited about the opportunity to delve more deeply into denominational differences in patterns of response, by race and class, and as they interact with region of country. In particular, we would like to apply what we know about the socioeconomic differences among different denominations to an understanding of denominational differences in response patterns. In future research, we will pursue with greater specificity the processes driving discrimination by race and class. Finally, we hope to use this project as a springboard to encourage future research involving stratification and religion.

OPPORTUNITIES AND PITFALLS OF IBFES FOR THE STUDY OF RELIGION

Field experiments are an under-utilized resource for religion scholars. In addition, IBFES offer numerous advantages for bringing stratification processes “back into” the study of religion. First, IBFES typically allow larger sample sizes, broader geographic coverage, and better representativeness of the underlying population than can be achieved through conventional field experiments. Larger sample sizes translate into greater statistical power and the use of large sample analytic techniques (Lahey & Beasley, 2009). Although we used a systematic sampling procedure to generate a national sample of churches, the method is adaptable to geographic units of any size – cities, counties, states, and even countries. Along similar lines, IBFES allow the researcher to increase the number of treatment conditions and move beyond the typical matched pairs methodology of traditional audit studies. For instance, in our church study, we had 4,680 churches spread across 6 races/class treatments and 12 denominations, meaning that we have 65 ($4,680 \div 72$) cases within each cell. This allows us to conduct statistically robust analyses on smaller segments of the file and to pursue rich comparisons among subgroups in the data set.

In addition, the costs of IBFES are relatively low. While we benefitted from a small army of undergraduate and graduate assistants who helped develop the sample list, send out e-mails, collect and sort the responses, code the data, and manage the data set, this does not differ significantly from the student-centered nature of many conventional field experiments (see Massey & Lundy, 2001). Considering that we contacted 4,680 churches nationwide, a traditional audit study on the same scale that requires auditors to appear in the field personally would have been infeasible. Similarly, while traditional correspondence studies can partially overcome the limitations of in-person field experiments, the costs of mailing, faxes, or other forms of correspondence would have added up quickly.

Also, the training and supervision requirements of IBFES are less demanding and the repercussions of mistakes by research assistants are not as severe. In-person field experiments typically require careful training of research assistants in order to insure uniform administration of experimental treatments. Similarly, in audit studies, auditors must be thoroughly trained as to how to present themselves in the field, must undergo periodic debriefing and retraining, and must guard against personal reactions when they experience discrimination. Similarly, although our research assistants

sometimes made mistakes – often due to our own poor instructions – these mistakes did not take place in the field where they could contaminate the results of the project. Moreover, because we exercised close supervision, mistakes were generally caught before they became major problems.

The use of IBFEs (and correspondence studies, in general) eliminates several potential sources of bias that might occur with in-person field experiments. Experimental researchers sometimes note that as experiments move from the laboratory to the field, there is a trade-off between internal validity and external validity. Because of the ability to rigorously control for alternative explanations, laboratory experiments have high internal validity but sometimes suffer from a lack of realism that hinders external validity. On the other hand, because field experiments take place in natural settings where subjects live, work, play, and pray, they have high external validity, but sometimes to the detriment of internal validity. Harrison and List (2004, p. 1033) question the necessity for this trade-off, noting: “An ideal field experiment not only increases external validity, but does so in a manner in which little internal validity is foregone.” IBFEs offer the potential to fully realize this goal. Because of larger sample sizes and random assignment of subjects (in our case, churches) to treatment conditions, we minimize the influence of extraneous variables that could pose as alternative explanations for the outcome variable.

Even so, researchers using IBFE methods must be aware of possible threats to internal validity. In our study, the biggest potential threats are inequalities due to place – at the macro-, meso-, and micro-levels. At the macro-level, regional/state variations might occur in the quantity or quality of church responses or how these interact with church denominations. Preliminary analyses suggest that macro differences in our sample are not large, probably because of random assignment. Meso-level differences occur, for instance, when churches sampled within the same district fall into communities with very different socio-geographic characteristics. Finally, micro-level differences occur when there are idiosyncratic differences between churches even when they fall in the same neighborhood. Although these potential threats to internal validity exist, the IBFE method provides a means for addressing them by geocoding data at the regional/state, zip code, and church levels. Our study is the first one we know of to incorporate higher level data such as these to measure attributes of the subjects (i.e., churches) as well as their context.

In addition, because many aspects of social life are increasingly experienced in online environments, the Internet provides access to more venues where inequality exists. For instance, a study like ours would have

been impossible 20 years ago before most churches had websites, e-mail, and online access. Similarly, the availability of online housing and rental sites and employment sites has opened up new possibilities for researchers interested in religious discrimination in housing or labor markets. Indeed, online access to a variety of organizational settings such as schools, nonprofit organizations, private business firms, social clubs and recreational centers, newspapers, government agencies, and universities represent untapped possibilities for IBFE-based research with a religious focus. Social networking sites and online chatrooms provide other venues for scholarly research on religion.

For all the advantages IBFEs offer, they are also “minimally invasive” (Harrison & List, 2004, p. 1035) in that they create very little disturbance to the natural setting. When people are contacted online, they can answer at their convenience or simply delete the request. In our case, we simply asked for basic information about their church, and so their responses, though highly informative, were retrieved at relatively low cost. Contrary to what some may expect, this type of research generally faces few difficulties in being cleared with Institutional Review Boards (IRBs) (Lahey & Beasley, 2009). Our IRB provided an expedited review of our proposal and simply required that students and faculty working on the project go through the campus IRB training.

As we have seen, the “data” derived from IBFE studies need not be limited to simple response rates as in most traditional correspondence studies. Rather, e-mails elicit textual material, which lends itself to content analysis and provide a rich trove of qualitative data. We received over 2700 e-mail responses in our church study. This type of data is not readily available in traditional audit studies or correspondence studies. Each level of the analysis provides a deeper level of understanding of the underlying processes creating discrimination. In addition, the widespread availability of websites makes accessible another source of data to learn about the organization being studied. In our study, churches routinely posted information about their staff and congregations, including photographs, allowing us to determine the race and gender of the head pastor as well as racial composition of members and church leaders. These data will be useful in making more robust determinations about discriminatory treatment.

In addition, the types of data retrieved from IBFE methods are expandable to suit the research purposes. For instance, in our church study we might have followed up two weeks after sending our first e-mail with those churches who *did not* respond, with a second request for information. For those churches that *did* respond, we might have sent a second e-mail

asking them for more information, like whether they could recommend a real estate agent in the community. These requests would surely have elicited another layer of data and enhanced our ability to detect discriminatory treatment. We even envision circumstances where IBFE methods could accommodate brief surveys using SurveyMonkey or some similar software.

Despite all these advantages, IBFEs also present challenges to the researcher. Perhaps the biggest challenge is that the relative novelty of the method and the myriad ways in which it might be deployed mean that there is no well-established track record of “do’s” and “don’ts” for researchers to follow. We often found ourselves inventing ways to address new issues that arose in the research. For instance, in seeking a way to generate a national sample, we came up with the idea of sampling Congressional districts arranged according to population density. Similarly, we discovered national websites for most of the denominations that aided our sampling within denominations (another resource that would not have been available before the Internet age). One of the biggest unanticipated issues was the amount of attention that was required to process (not simply code) the church responses as they came in. A higher percentage than we had thought were ambiguous responses that required special attention. For instance, we received some automated responses and we had to decide how to handle these. We decided to count them as no response. In other cases, our initial e-mails were received by an administrative office that represented multiple churches. In these cases, we either followed up with one of the member churches or dropped that case from the analysis.

In addition, the day-to-day administration of the project took a heavier toll than we expected. First, as with many research projects, ours required a lot of planning and preparation up front, but with so few previous models for us to follow, the planning phase took up more time than usual. Second, the project was very labor-intensive and required the assistance of over a dozen undergraduate and graduate assistants at various phases of the project. The project alternated between bursts of intensive activity (as when the initial e-mails were being sent out) and long periods of monotonous, but essential, grunt work (such as the content analysis of the churches’ responses). These various phases required constant supervision and weekly meetings among members of the research team. Third, the multiple phases of the project presented special data management challenges. Our project integrates five large data sets – the master data set with all the church contact information and coded responses, the content analysis data set, the church website data set, the zip code data set, and the qualitative data – which had to be accurately matched with identification codes. For the

content analysis phase, we had to replace the master ID number for each case with “super-secret ID numbers” so that coders could not identify the class or race of the character.

Perhaps the main limitation of IBFEs is that only a relatively narrow slice of social experience is accessible by Internet-based observation or manipulation. Internet manipulations tend to be one dimensional – in our case, written e-mails – as opposed to richer penetrations into the field afforded by in-person field experiments. Of course, as we suggested above, IBFEs need not be limited to a single contact, but even so the range of experience they can access is limited. Also, in-person experiments provide opportunities to pick up important cues from the social context that aid in the interpretation of the data. We can only partially compensate for this loss of contextual data with the church website and zip code data. There is a lot that sociologists of religion would like to know that cannot be easily accessed by an e-mail or website. In our own research, we would be interested in knowing the experiences of different types of people as they visit churches for the first time, but this topic is beyond the reach of online experiments.

A perpetual concern about IBFEs (and correspondence studies in general) is the nature of the manipulation, particularly whether it is too weak to be noticed by the subjects or perhaps too strong.⁵ In our case, the key manipulation was the use of names that would likely be associated with persons of certain races or ethnicities by most people. How might our results have changed if the e-mails had been accompanied by a family picture that reinforced those racial identities, but also humanized the characters sending the e-mails? A secondary manipulation in our study was the tag lines added to some of the white characters’ e-mails signifying their class status as poor or rich. These manipulations were deliberately subtle, almost understated, and may have been so weak to contribute in our analyses to lower class differences than race differences. A stronger manipulation – perhaps a follow-up sentence elaborating on the class situation – might have elicited sharper class differences. Another concern about the manipulation is that it will be perceived in a way that was not intended, as with one respondent who interpreted an e-mail from our poor white character Greg Murphy as “just a spam letter to get information.” We have no way of knowing whether other recipients of our e-mails might have reacted similarly to the “lost my job” or “sold my company” manipulation.

Two related concerns of IBFEs articulated by [Lahey and Beasley \(2009\)](#) are template bias and experimenter bias. Template bias occurs when the researcher provides a limited number of templates for each manipulation.

Experimenter bias is introduced when the researcher is forced to make choices among limited templates. In our study, we essentially used 12 templates – 6 different race/class manipulations \times 2 names per manipulation – or only 2 templates per manipulation. Lahey and Beasley (2009) would argue, with some justification, that our results would have been more robust had we used multiple templates of the initial e-mail and multiple names for each race or ethnicity. A full implementation of Lahey and Beasley's suggestions would assure robustness, but would also complicate the research design and increase the administrative effort required. Also, the number of templates per treatment might be limited in designs with more treatment conditions. In any event, Lahey and Beasley (2009) offer a new computer program that will assist in implementing more complicated template designs.

Bottom line, we draw several lessons from our early experiences with IBFEs. First, IBFEs are administratively more complex than meets the eye and should not be undertaken lightly. It is important that researchers using this method engage in careful planning and preparation, maintain close supervision of the project, be able to adapt to unanticipated circumstances, and "sweat the details." Second, for some time to come, we will likely be conducting such studies in relatively uncharted waters, particularly as IBFEs penetrate new social environments and develop more complicated research designs. This requires that researchers using this method share their knowledge and continue to learn from each other's mistakes. Third, due to the relatively small base of experience from which we have to draw, a higher percentage of decisions than we would like to admit are "trial and error." In order for the learning curve for IBFE researchers to progress, we must document our experiences with this method with the ultimate goal of developing a protocol of best practices for handling recurring problems and situations. Despite these caveats, IBFEs offer many advantages that make them an important addition to the methodological toolkit of sociologists of religion and stratification.

CONCLUSIONS

In this chapter, we have extolled the virtues of IBFEs as an innovative, new method that can help in bringing stratification processes "back into" the study of religion. Religious scholars in sociology have been surprisingly remiss in failing to integrate conventional field experiments into their methodological repertoire. They should not continue to be laggards as field experiments move into the Internet age.

IBFEs offer religion scholars rich opportunities to investigate stratification processes in general, and discriminatory treatment in particular. Two major genres of discrimination studies can be envisioned. First, as in this study, researchers can ask, “How do different religious traditions discriminate by race, class, age, education level, sexual orientation, immigrant status, disability, political identification, or any other status characteristics?” A second question is, “To what extent does intensity of religious identification or religious tradition (i.e., Catholic, Protestant, Jew, Muslim, Pagan, Atheist, etc.) as a status characteristic result in discriminatory treatment in the job market, the housing market, or consumer settings?” The possibilities of addressing broader questions of stratification, such as the similarities and differences among different religious traditions or denominations by intensive examination of church websites, are also possible.

In many respects, the Internet holds the potential to revolutionize field experiments the way that the telephone revolutionized survey research methodology. Just as telephone surveys improved upon face-to-face surveys by extending the geographic reach, the potential sample size, and the representativeness of survey research at lower cost, IBFEs promise to do the same with the field experiment method. Telephone surveys dramatically reduced (but did not eliminate) sources of interviewer bias; IBFEs accomplish even better diminution of bias compared to in-person field experiments. Just as the advent of telephone surveys created new, previously unimaginable opportunities as the methodology evolved and penetrated different venues of social life, IBFEs will likely open up new possibilities for scholars of religion and other areas. Also, telephone survey methodology initially faced a steep learning curve and lurched forward by trial and error, but with the passage of time its practitioners crafted a crude understanding of the “do’s” and “don’ts” that eventually morphed into a highly codified system of best practices that today enjoy a broad consensus in the field. Is there any reason to believe that IBFEs cannot do the same?

IBFEs hold great promise for the field of sociology. Religion scholars should not be left behind in this exciting new venture.

NOTES

1. The earliest Internet-based field experiment we could find on any topic is by [Lucking-Reiley \(1999\)](#), who studied the effects of different auctioning techniques for the sale of game cards on the Internet.

2. One of the variables content coded was the size of the church. Since the original e-mails inquired about the size of the congregation, most respondents indicated

church size in terms of either number of church members, weekly attendance at services, or both.

3. Multiple responses took two forms: (a) responses from more than one church official from the same church and (b) multiple responses from the same respondent.

4. The U.S. Bureau of the Census divides the country's geography into four "regions" and nine "divisions." In this chapter, we use the term "regions" to identify the nine units that the Census defines as "divisions."

5. Ideally experiments conduct manipulation checks, which allow researcher to test the strength of the manipulation. With IBFEs, however, this can often be difficult to accomplish without extensive pretesting.

ACKNOWLEDGMENT

This chapter was presented at the Bringing Stratification Processes "Back In" to the Scientific Study of Religion: A Penn State Stratification and Social Change Conference, University Park, PA, May 20–21, 2011. We thank participants at the conference for comments on an earlier version of the chapter.

REFERENCES

- Ahmed, A. M., Andersson, L., & Hammarstedt, M. (2008). Are lesbians discriminated against in the rental housing market? Evidence from a correspondence testing experiment. *Journal of Housing Economics, 17*, 234–238.
- Ahmed, A. M., & Hammarstedt, M. (2009). Detecting discrimination against homosexuals: Evidence from a field experiment on the Internet. *Economica, 76*, 588–597.
- Banerjee, A., Bertrand, M., Datta, S., & Mullainathan, S. (2009). Labor market discrimination in Delhi: Evidence from a field experiment. *Journal of Comparative Economics, 37*, 14.
- Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review, 94*, 991–1013.
- Carey, R. G. (1971). Influence of peers in shaping religious behavior. *Journal for the Scientific Study of Religion, 10*, 157–159.
- Carpusor, A. G., & Loges, W. E. (2006). Rental discrimination and ethnicity in names. *Journal of Applied Social Psychology, 36*, 934–952.
- Correll, S. J., Benard, S., & Paik, I. (2007). Getting a job: Is there a motherhood penalty? *American Journal of Sociology, 112*, 1297–1339.
- Dovidio, J. F., Gluszek, A., John, M.-S., Dittmann, R., & Lagunes, P. (2010). Dimensions of difference and experience of exclusion. *Journal of Social Issues, 66*, 59–78.
- Edwards, K. L. (2008). Bringing race to the center: The importance of race in racially diverse religious organizations. *Journal for the Scientific Study of Religion, 47*, 5–9.
- Festinger, L., & Katz, D. (Eds.). (1953). *Research methods in the behavioral sciences*. New York, NY: Dryden Press.

- Harrison, G. W., & List, J. A. (2004). Field experiments. *Journal of Economic Literature*, 42, 1009–1055.
- Heckman, J. J. (1998). Detecting discrimination. *Journal of Economic Perspectives*, 12, 101–116.
- Johnson, W. T. (1971). The religious crusade: Revival or ritual? *American Journal of Sociology*, 76, 873–890.
- Lahey, J. N., & Beasley, R. A. (2009). Computerizing audit studies. *Journal of Economic Behavior and Organization*, 70, 508–514.
- LaPiere, R. T. (1934). Attitudes versus actions. *Social Forces*, 13, 230–237.
- Lucking-Reiley, D. (1999). Using field experiments to test equivalence between auction formats: Magic on the Internet. *American Economic Review*, 89, 1063–1080.
- Malhotra, D. (2010). (When) are religious people nicer? Religious salience and the “Sunday effect” on pro-social behavior. *Judgment and Decision-Making*, 5, 138–143.
- Massey, D. S., & Lundy, G. (2001). Use of black English and racial discrimination in urban housing markets: New methods and findings. *Urban Affairs Review*, 36, 452–469.
- McCloud, S. (2007). Putting some class into religious studies: Resurrecting an important concept. *Journal of the American Academy of Religion*, 75, 840–862.
- Pager, D. (2003). The mark of a criminal record. *American Journal of Sociology*, 108, 937–975.
- Pager, D. (2007). The use of field experiments for studies of employment discrimination: Contributions, critiques, and directions for the future. *Annals of the American Academy of Political and Social Science*, 609, 104–133.
- Pager, D., & Quillian, L. (2005). Walking the talk: What employers say versus what they do. *American Sociological Review*, 70, 355–380.
- Pager, D., Western, B., & Bonikowski, B. (2009). Discrimination in a low wage labor market: A field experiment. *American Sociological Review*, 74, 777–799.
- Paluck, E. L. (2010). The promising integration of qualitative methods and field experiments. *Annals of the American Academy of Political and Social Science*, 628, 59–71.
- Riach, P. A., & Rich, J. (2006). An experimental investigation of sexual discrimination in hiring in the English labor market. *Advances in Economic Analysis and Policy*, 6, 1–20.
- Rooth, D.-O. (2009). Obesity, attractiveness, and differential treatment in hiring: A field experiment. *Journal of Human Resources*, 44, 710–735.
- Smith, C., & Faris, R. (2005). Socioeconomic inequality in the American religious system: An update and assessment. *Journal for the Scientific Study of Religion*, 44, 95–104.
- Soetevent, A. R. (2005). Anonymity in giving in a natural context – A field experiment in 30 churches. *Journal of Public Economics*, 89, 2301–2323.