There has been a good deal of experimental research reported in the past three decades about the nature of the mailings that can be sent to sampled respondents, such as in surveys that use an addressed-based sampling frame. This literature has consistently shown that there are many features of such mailings that can have positive effects on survey response rates, but the literature has almost entirely focused on the attributes of what is enclosed inside the envelope (e.g. the content of the letter, whether a cash incentive is enclosed, and other materials that may be enclosed). In contrast, very little attention appears to have been given to the importance of the envelope/mailer that is used to send the survey communication. However, logic dictates that whatever is enclosed inside the envelope can have no impact on the recipient unless the envelope is opened. Past experience also has shown that Hispanics are consistently underrepresented in sample surveys in the United States, and this is especially true for Spanish-dominant Hispanics. To that end, we report on a large national experiment (n=9,919) that tested the effects of what was used for the first line of the envelope’s address in a mailing sent to randomly sampled addresses in high density Hispanic neighborhoods in the United States. This experiment tested the effects on survey response of whether the envelope was addressed one of three ways: Current Resident or Current Resident/Residente Actual or Residente Actual/Current Resident. Overall, it was found that using Residente Actual/Current Resident as the addressee led to (1) higher return rates of completed questionnaires and (2) a greater portion of returned completed Spanish-language questionnaires. In addition, it also was found that meaningful regional differences existed as to which addressee line generated the most positive response.

INTRODUCTION

The research study that is presented in this paper is based on a national experiment that investigated the effects of what was written in the addressee line of the envelope used in a mail survey of Hispanics. The paper starts with brief literature reviews of what is known about (1) the features of what is mailed to sampled respondents in mail surveys and (2) the challenge of gaining survey cooperation from Hispanics, especially those who are Spanish-dominant.
MAIL SURVEY ENVELOPES
There is a fair amount of past research on the effects of the nature of the mailings in surveys where the address of sample households are known, but this body of literature is neither exhaustive nor does it provide definitive knowledge about how to deploy mailings to sampled households so as to achieve the most cost-effective outcomes (e.g. Camburn et al. 1996; de Leeuw et al. 2007; Dillman et al. 2009; Groves and Snowden 1987). In reviewing this literature, there appear to be two major domains of factors associated with the mailings in mail surveys that have been tested (or could be tested) to raise survey response:

1. Factors related to the envelope and the mailing service used to deliver it, such as the postage used on the envelope, the sender that is listed in the return address, graphics on the envelope, or the size of the envelope. Figure 1 shows factors that fall into this domain.

2. Factors related to the substance and formatting of the materials sent inside the mailing itself (e.g. the substantive information conveyed inside the envelope about the purpose and confidentiality of the study, the letterhead on which a cover letter is printed, font size and font type of the text, and the credentials of the sender and signatory), and what else may be included inside the envelope (e.g. a noncontingent incentive, mention of a contingent incentive, an informational/motivational brochure, or an FAQ insert).
The first domain of factors (Figure 1) is very important because the contents of the mailing (including the possibility of finding a noncontingent incentive inside the envelope) cannot have any effect on survey response unless the mailer is opened. Thus, the first goals in trying to successfully deploy survey mailings are to (1) get the envelope delivered to the correct address and (2) get people to open the envelope that is delivered to them. Little has been reported about factors that are related to the envelope and the mail delivery service used to deliver it and how these factors affect the propensity of a recipient opening it. This dearth of knowledge is readily apparent when one considers what de Leeuw et al. (2007) did not report in their otherwise extensive meta-evaluation of the effects of advanced letters; i.e. they included no factors from this envelope-domain in their meta-evaluation.

In surveys in which the household address is the sampling unit, it normally does not matter to the researchers exactly who is/are the current resident(s) of the household. Thus, the name of the current resident(s), where the name can be matched to the address, often is not used in the addressee line of the mailing. The reasoning for this is that it is not always the named person from whom the survey technically is trying to gather data, but rather from whoever currently

<table>
<thead>
<tr>
<th>Factors</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of mailer and type of delivery service/class</td>
<td>USPS 1st Class, USPS Priority Mail, Fed-Ex, UPS</td>
</tr>
<tr>
<td>Envelope size</td>
<td>Small letter, standard letter, 6&quot; x 9&quot;, 9&quot; x 12&quot;</td>
</tr>
<tr>
<td>Envelope color</td>
<td>White; business tan; pastel</td>
</tr>
<tr>
<td>Addressee on envelope</td>
<td>Name of resident, current resident, consumer, head of household</td>
</tr>
<tr>
<td>Return address of sender</td>
<td>University, government agency, survey organization, corporate sponsor</td>
</tr>
<tr>
<td>Graphics on envelope</td>
<td>Cartoon art, scenic image, celebrity picture</td>
</tr>
<tr>
<td>Branding logo on envelope</td>
<td>Organizational branding, branding of the survey</td>
</tr>
<tr>
<td>Topical phrasing on envelope</td>
<td>“Childhood vaccinations,” “Crime prevention awareness,” “Educational attainment,” “U.S. foreign policy”</td>
</tr>
<tr>
<td>Endorsement on envelope</td>
<td>Elected official, famous athlete, music celebrity, TV celebrity</td>
</tr>
<tr>
<td>Arrival of letter compared to start of data collection</td>
<td>Number of days the envelope arrives prior to contact for data collection</td>
</tr>
<tr>
<td>Font style</td>
<td>Times-Roman; Arial; Courier; Verdana; Calibri</td>
</tr>
<tr>
<td>Font size</td>
<td>10, 11, 12, 14</td>
</tr>
</tbody>
</table>

*Figure 1* Factors related to the likelihood the mailer will be opened and acted upon.
lives there. By listing a name of someone thought to live at an address in the addressee line, the researchers likely will cause (1) the mailing to be forwarded to another (ineligible) address if the named person has moved and (2) the current resident(s) to throw away the mailing if it is addressed to someone else who no longer lives there. That is why many such mail surveys address their envelopes to Current Resident, Current Householder, or some such generic nomenclature, sometimes along with the matched name for the address, e.g. Mary Jones or Current Resident.

However, doing so runs contrary to four decades of advice from Don Dillman and his colleagues (e.g. Dillman et al. 2009) to do as much as possible to tailor mailings to appeal to the recipient. But in the case of what to use for the addressee line on a mailing envelope in a survey of the general population, how can a researcher do anything to tailor it beyond using some generic wording? Well, one thing that can be done is to try to convey something in the addressee line that will have special appeal to the sampled persons from whom the survey is trying to gain cooperation.

To that end, Lavrakas et al. (2015) recently have reported success by tailoring the addressee line of an envelope. In the Lavrakas et al. study, a survey was to be conducted of a relatively small subset of the general public with the eligible sample having to be identified within a much larger sample of the general population. This survey was about the vaccination history of toddlers and other very young children, and it was their parents/guardians who were the eligible respondents for the survey. In their experimental study, these researchers found that envelopes address to Parent/Guardian led to markedly and significantly better response rates among eligible households than did envelopes addressed to Current Resident or to the name of a person matched to the address of the mailing. These findings led Lavrakas et al. to opine:

- This suggests that researchers studying other topics and other populations of respondents will need to think carefully about whether and how best to tailor the addressee line to their particular study. For example, in a survey about ... the political preferences of those without any party affiliation using [an addressed based] sampling frame, one could envision an advance letter addressed to ‘Independent-minded Voter.’ (p. 19)

**GAINING COOPERATION FROM HISPANICS IN SURVEYS IN THE UNITED STATES**

It has long been known that Hispanics in the United States are markedly less likely to respond to survey requests than are non-Hispanics. This is especially true among so-called Spanish-dominant Hispanics who are defined as those who use Spanish as their only language or their primary language within their household. How this is known to happen beyond any doubt in almost all general population surveys in the United States is that these surveys all must
upweight the Hispanics from whom they do gather data before performing their analyses, because they consistently undersample Hispanics, often to the extent of gathering data from less than half of the number of Hispanics that should have had been interviewed for the survey (cf., Lavrakas et al. 2011).

Furthermore, in upweighting the Hispanics in their samples from whom data have been gathered, researchers in the United States most often fail to have an accurate “mix” of Hispanics to use in their weighting. That is, not only do they underrepresent Hispanics in their final unweighted samples, but the Hispanics from whom they do gather data often are woefully unrepresentative of the Hispanic population in the United States because the vast majorities from whom data have been gathered are English-dominant Hispanics. This compares to population parameters that Nielsen generates with its very high quality National Hispanic Enumeration Survey; this survey indicates that nearly half of all Hispanics in the United States are Spanish-dominant. As such, far too few Spanish-dominant Hispanics have data gathered from them in most gen-pop surveys in the United States, even in those surveys for which the data can be gathered in Spanish. (In fact, it is not unusual for one-fifth or less of the number of Spanish-dominant Hispanics that should have been interviewed actually does provide data for most surveys.) And, because Spanish-dominant Hispanics are very different from English-dominant Hispanics in terms of their other demographics, psychographics, life styles, attitudes, and preferences (cf., Lavrakas et al. 2011), most surveys conducted in the United States fail to accurately measure these statistics for the Hispanic population living in the United States.

Therefore, it is important for survey researchers to find new and cost-effective methods to achieve more success in gathering data from U.S. Hispanics in their surveys, and in particular from Spanish-dominant Hispanics.

PURPOSE OF THE PRESENT STUDY

In 2015, an experiment was conducted by Simmons Research to determine whether they could achieve more cost-efficiencies in recruiting Hispanic households into their National Hispanic Consumer Study (NHCS). This experiment addressed two goals: (1) increase success in gaining cooperation in the NHCS and (2) increase success in gaining cooperation from Spanish-dominant Hispanics.

In this experiment, randomly sampled addresses in areas with relative high Hispanic density were randomly assigned to one of three experimental conditions that determined what was used as the addressee line in the mailing that was done to the households asking them to complete an enclosed NHCS enumeration questionnaire and then mail it back to Simmons. There were two general hypotheses that were tested in this experiment:

1. The content of the addressee line used on the outbound mailing envelope from Simmons would affect the rate of returned completed
METHODS

BACKGROUND ON THE NHCS

The Simmons NHCS deploys a probabilistic addressed-based sampling design to produce representative measures of consumer behavior and attitudes to products, brands, and media among all Hispanics/Latinos in the United States—be they English-speaking or/and Spanish-speaking. The NHCS uses a two-phase data collection approach, with Phase 1 consisting of mail-based recruitment and an enumeration questionnaire to obtain the household’s agreement to participate in the study and to gather basic media exposure and demographic data about the household. The questionnaire is sent in both Spanish and English. Phase 2 involves the mailing of self-administered survey booklets to eligible household members (in the language most appropriate for each) to gather very detailed information about their consumer behaviors, attitudes, and preferences.

EXPERIMENTAL DESIGN

For the experiment, sampled addresses were randomly assigned to one of three addressee line conditions:

- Control condition: Current Resident
- Treatment condition 1: Current Resident/Residente Actual
- Treatment condition 2: Residente Actual/Current Resident

Traditionally, Simmons has used “Current Resident” as the addressee line in its NHCS mailings. Therefore, this served as the Control Condition in this experiment. It was hypothesized that adding the Spanish equivalent of “Current Resident,” which is “Residente Actual,” would (1) help differentiate the mailing from other unsolicited incoming mail and (2) have special appeal to Hispanics, especially those who are Spanish-dominant.

SAMPLING ISSUES

A total of 9,919 randomly sampled addresses were used in this experiment and were randomly assigned to the three addressee line conditions as follows:

- Current Resident, n=3,307
- Current Resident/Residente Actual, n=3,306
- Residente Actual/Current Resident, n=3,306
The success of this random assignment was confirmed by conducting a series of analyses using frame data to determine whether the three groups were in fact equivalent to each other, and they were found to be so.

Furthermore, there were two types of samples used in this experiment. Once was addresses in high density Hispanic areas that were randomly sampled for the first time and had never been contacted by Simmons prior to this study (n=6,447). The other were nonresponding addresses in a previous NCHS (n=2,789).

**DEPENDENT MEASURES**

There were two primary dependent variables of interest, each of which was a binary (0/1) variable:

- **Returned Questionnaire:** whether a completed NHCS enumeration questionnaire was returned to Simmons (Yes=1; No=0)
- **Language of Returned Questionnaire:** whether the English or Spanish questionnaire was completed and returned (Spanish=1; English=0)

**ADDITIONAL DATA APPENDED TO THE FRAME**

To allow the researchers to conduct more precise and informative analyses of the experimental effects, data was appended to each of the 9,919 addresses used in our experiment. These auxiliary data were measures of Census characteristics for the local zip code within which each address was located. The seven variables that we appended were chosen because of our past experience that each was predictive of response/nonresponse to survey recruitment efforts: (1) Pct. Population, Black/African American Alone; (2) Pct. Population, Hispanic/Latino; (3) Pct. Housing Units, Renter-Occupied; (4) Pct. Housing Units, Mobile Home; (5) Pct. Population 16+, Female; (6) Civilian Labor Force, Employed; Pct. Population 25+, Non Graduated High School; and (7) Pct. Population 25+, College Graduates.

**RESULTS**

**COMPLETION RATES BY EXPERIMENTAL CONDITION**

**Rate of Return of Completed Questionnaires**

As shown in Table 1 in the Total Sample, there was a statistically significant difference ($\chi^2 [2]=7.9$, $p<0.02$) between the rates at which completed questionnaires were returned and what was listed in the addressee line on the envelope. Residente Actual/Current Resident had the highest response rate and did 14 percent (2.7pp) better than the control condition (Current Resident). The Residente Actual/Current Resident also had the highest response rate within the New Sample and within the Previous Nonresponder Sample; that was 13 percent (3.0pp) higher than the control condition in the New Sample, and 17 percent (2.0pp) higher than the control condition in the
previous Nonresponder Sample.

<table>
<thead>
<tr>
<th></th>
<th>Current resident</th>
<th>Current resident/residente actual</th>
<th>Residente actual/current resident</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pct. Complete from Total Sample**</td>
<td>19.8% (n=3,125)</td>
<td>20.3% (n=3,138)</td>
<td>22.5% (n=3,163)</td>
</tr>
<tr>
<td>Pct. Complete from New Sample*</td>
<td>23.3% (n=2,209)</td>
<td>24.1% (n=2,211)</td>
<td>26.3% (n=2,227)</td>
</tr>
<tr>
<td>Pct. Complete from Previous Nonresponder Sample</td>
<td>11.6% (n=926)</td>
<td>11.2% (n=927)</td>
<td>13.6% (n=936)</td>
</tr>
</tbody>
</table>

*p<0.10; **p<0.05.

In addition to these overall results, the effect of the addressee line varied by region of the country in which a sampled address was located. In the East North Central region, it was Current Resident that elicited the highest regional return rate of completed questionnaires. In the Middle Atlantic, South Atlantic, East South Central, and Mountain regions, it was Current Resident/Residente Actual that elicited the highest regional return rate of completed questionnaires. Whereas in the New England, West South Central, West North Central, and Pacific regions, it was Residente Actual/Current Resident that elicited the highest regional return rate of completed questionnaires. Of note, only in the West South Central region and the Pacific region are the observed differences by Census region statistically significant. But it also is important to note that it is in these regions where a large number of Hispanic residents of the United States live.

**Rate of Return of Completed Spanish Questionnaires**

Table 2 presents the proportion of returned completed questionnaires that were returned using the Spanish language version of the questionnaire. Overall, the Residente Actual/Current Resident addressee line had the highest Spanish questionnaire return rate and did 4 percent (1.1pp) better that the control condition. The Residente Actual/Current Resident also had the highest Spanish return rate within the New Sample; that was 3 percent (0.8pp) higher than the control condition. Within the sample of Previous Nonresponders, the Current Resident/Residente Actual addressee line had the highest rate of returned Spanish questionnaire, which was 15 percent (2.7pp) higher than the control condition. However, none of these differences achieved even marginal levels of significant; all were p>0.10.
The analytic comparisons that are presented in Tables 1 and 2 were also tested using ANACOVA (Analysis of Covariance) with a number of census characteristics of local neighborhoods and the design-related variables (e.g. incentive levels) used in the study serving as covariates. These F-tests led to the same conclusions about effect sizes and statistical significance across the experimental conditions as did the chi-square tests reported above.

**LOCAL AREA CHARACTERISTICS AND RETURN RATES**

As shown in Table 3, there are several significant correlations between local area census characteristics, at the zip code level, and (1a) whether or not a type of addressee line led to a completed returned questionnaire and/or (2) whether the questionnaire that was returned was the Spanish-language one.

---

Table 2  Overall return rate of completed Spanish questionnaires from among all completed questionnaires by addressee line.

<table>
<thead>
<tr>
<th></th>
<th>Current resident</th>
<th>Current resident/ residente actual</th>
<th>Residente actual/ current resident</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pct. Spanish from total sample**</td>
<td>26.1% (n=621)</td>
<td>26.4% (n=637)</td>
<td>27.2% (n=712)</td>
</tr>
<tr>
<td>Pct. Spanish from new sample*</td>
<td>27.8% (n=514)</td>
<td>27.4% (n=533)</td>
<td>28.7% (n=585)</td>
</tr>
<tr>
<td>Pct. Spanish from previous nonresponder sample</td>
<td>17.8% (n=107)</td>
<td>21.2% (104)</td>
<td>20.5% (127)</td>
</tr>
</tbody>
</table>

*p<0.10; **p<0.05.
## Table 3  Correlations between local area Census characteristics and response (n=9,896).

<table>
<thead>
<tr>
<th>Zip code Census characteristics</th>
<th>Return of completed questionnaire by type of addressee line</th>
<th>Return of Spanish questionnaire by type of addressee line</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Current resident resident/ current resident</td>
<td>Current resident resident/ current resident</td>
</tr>
<tr>
<td>Pct. population, Black/African American alone (Mean=12.1; SD=14.2)</td>
<td>-0.037*</td>
<td></td>
</tr>
<tr>
<td>Pct. population, Hispanic/Latino (Mean=44.1; SD=26.4)</td>
<td>0.263***</td>
<td>0.182***</td>
</tr>
<tr>
<td>Pct. housing units, renter-occupied (Mean=38.0; SD=16.9)</td>
<td>-0.041*</td>
<td>-0.054**</td>
</tr>
<tr>
<td>Pct. housing units, mobile home (Mean=3.7; SD=6.3)</td>
<td>0.048**</td>
<td></td>
</tr>
<tr>
<td>Pct. population 16+, female: civilian labor force, employed (Mean=41.7; SD=6.7)</td>
<td>-0.125***</td>
<td></td>
</tr>
<tr>
<td>Pct. Population 25+, Non Graduated High School (Mean=20.3; SD=12.7)</td>
<td>-0.042*</td>
<td>-0.061**</td>
</tr>
<tr>
<td>Pct. Population 25+, College Graduates (Mean=33.6; SD=16.3)</td>
<td>0.040*</td>
<td>-0.042*</td>
</tr>
<tr>
<td>(n)</td>
<td>3128</td>
<td>3131</td>
</tr>
</tbody>
</table>

*p<0.05; **p<0.01; ***p<0.001.
The correlations shown in Table 3 are only those that exceeded the 0.05 level of significance, but with these relative large sample sizes even small correlations will be significant.

In terms of the local area characteristics of households, there was some variation in which characteristics were predictive of a returned completed questionnaire among the type of addressee line that was used on the mailing envelope sent to the sampled household address:

- A Current Resident addressee was less likely to return a complete questionnaire in zip codes with a greater percentage of residents not graduating from high school.
- A Current Resident/Residente Actual addressee was (1) less likely to return a complete questionnaire in zip codes with a greater percentage of residents not graduating from high school and (2) in zip codes with a greater percentage of renter-occupied households, but (3) was more likely to return a complete questionnaire in zip codes with a greater percentage of residents graduating from college.
- A Residente Actual/Current Resident addressee was (1) less likely to return a complete questionnaire in zip codes with a greater percentage of residents not graduating from high school, (2) in zip codes with a greater percentage of Black residents, and (3) in zip codes with a greater percentage of renter-occupied households, but (4) was more likely to return a complete questionnaire in zip codes with a greater percentage of residents living in a mobile home.

The size of the significant correlations found between zip code level census characteristics and whether a Spanish language questionnaire was returned were noticeably stronger than the correlations discussed above. In terms of the local area characteristics of households, there was little variation in which characteristics were predictive of a returned completed Spanish-language questionnaire by the type of addressee line that was used on the mailing envelope sent to the sampled household address:

- Current Resident addressees and Residente Actual/Current Resident addressees were (1) less likely to return a Spanish questionnaire in zip codes with a greater percentage of residents graduating from college and (2) in zip codes with a greater percentage of female residents employed, but (3) were more likely to return a Spanish questionnaire in zip codes with a greater percentage of Hispanic residents, (4) renter-occupied households, and (4) greater percentage of resident not graduating from high school.
- A Current Resident/Residente Actual addressee was (1) more likely to return a complete questionnaire in zip codes with a greater percentage Hispanic residents, (2) residents not graduating from
high school, and (3) with a greater percentage of renter-occupied households, but (4) were less likely to return a Spanish questionnaire in zip codes with a greater percentage of residents living in a mobile home.

**DISCUSSION**

Although there has been considerable research on the topic of how to best gain cooperation in mail surveys, one topic that has not received much attention is how the features of the envelope that is used in the mail survey affect cooperation. If the envelope is not opened by the sampled recipient, then there is no chance for cooperation with the survey task. So a challenged faced by mail surveys is to differentiate a survey-related mailing that is sent to an address from the junk mail sent to that address and to motivate the recipient to open the envelope.

One way to succeed in doing this appears to be through the use of an appealing addressee line, especially when the survey is targeted to a subsection of the general population, e.g. Hispanics.

In the experiment that we conducted, we found the following findings for a national mail survey of Hispanic households:

- Overall, adding Residente Actual to the addressee line of our mail survey envelope raised the response rate above using only Current Resident.
- Overall, adding Residente Actual to the addressee line of our mail survey envelope raised the proportion of Spanish-language questionnaires among the returned completed questionnaires.
- Overall, the addressee line Residente Actual/Current Resident raised response rates more than Current Resident/Residente Actual.
- Overall, adding Residente Actual and especially in the case of the order, Residente Actual/Current Resident, yielded the highest response rates among addresses with a matched telephone number.
- Overall, the addressee line Residente Actual/Current Resident raised the proportion of Spanish-language questionnaires among the returned completed questionnaires more so than Current Resident/Residente Actual, except among the sample of previous nonresponders.
- There were meaningful regional differences in terms of which addressee line yielded the highest response rate.
- There are several local area Census characteristics that were significantly related to the rate of response for the different addressee lines.
• There are several local area Census characteristics that were significantly related to the proportion of Spanish-language questionnaires among the returned completed questionnaires for the different addressee lines.

In addition to these findings, it is worth noting that although the positive effects that we observed in the experiment were not enormous ones, they come at no financial cost to a research organization. That is, survey costs associated with using Residente Actual in the addressee line on an envelope are the same as the costs of using Current Resident or some other such designation. Furthermore, total survey costs will be lessened when an addressee line is more effective in yielding higher response rates.

These promising findings suggest that more research on the topic of the addressee line and on other topics related to the envelopes used for mail surveys is called for.
REFERENCES


