Measurement Bias: The Value of a Second Opinion

A White Paper By Edward T. Rincon, Ph.D.

Introduction

Despite their obvious value, there is a natural reluctance to seek second opinions, especially when important decisions are under consideration. Second opinions can cause one to question what is currently known about the world around us and lead to unplanned or undesired changes. In some instances, second opinions can lead to improvements as well.

As a case in point, generations of college aspirants have been influenced in some way by college admissions tests like the SAT, LSAT or GMAT. If it were not for the "second opinions" provided by grade point averages and recommendations, perhaps the careers of many college aspirants may have taken a different path. And who can forget the polling machinery that cursed the presidential election of 2000? Indeed, the election required a "second opinion" by the U.S. Supreme Court to resolve.

In the media and entertainment industry, the ratings provided by Nielsen Media Research greatly influence what we see on television. Unlike other high-stakes measurements, however, there are no "second opinions" in the television industry, a fact that has caused great concern when the Nielsen ratings lead to undesirable or controversial outcomes. Recent attention by the National Latino Media Council (NLMC), for example, has focused on the consequences of under-counting Latino television audiences. When Latino television audiences are under-counted, it leads to:

- The pre-mature cancellation of new programs
- The exclusion of Latino actors from television shows and movies
- Reluctance among network executives to invest in future Latino-themed shows; and
- The networks' loss of millions of advertising dollars from advertisers who want to target Latino audiences.

Indeed, the failure to show sufficient Latino audiences proved fatal for such Latino-themed programs as *Luis*, *Greetings from Tucson*, *Resurrection Boulevard*, and *Kingpin*. The one notable exception has been *The George Lopez Show*, which, by some measures, has succeeded in appealing to a diverse television audience.

These concerns have set in motion a series of events that would begin to challenge the Nielsen formula for measuring Latino television audiences.

Nielsen Comes Under Siege

Continued concern for the negative consequences posed by the Nielsen ratings fueled the drive to challenge Nielsen Media Research along several fronts. First, the National Latino Media Council (NLMC) launched the Latino Television Study in August of 2003 to explore the potential under-count of Latino television audiences. This study, funded by an unconditional grant from CBS and ABC, pointed to a potentially large under-count of Latinos viewing The George Lopez Show during the Fall of 2003.

Secondly, in February of 2004 Nielsen Media Research released the controversial results of its local people meter (LPM) test in the New York market, which revealed sharp, unexplained drops in viewership (ranging from 25 to 60 percent) for shows popular with Latino and African-American audiences. Although Nielsen staff explained that the differences were expected because the LPM was a more accurate measuring device than the traditional written diaries, various advocacy groups remained skeptical about the accuracy of the LPM. One reason for the skepticism stemmed from a previous incident wherein Nielsen Media Research admitted to an under-count of 300,000 Latino viewers in the New York market. [1]

Thirdly, growing concern about the potential under-count led to the creation of the *Don't Count Us Out Coalition*, a group consisting of leading community and grassroots leaders who were driven to raise public awareness of Nielsen Media Research's systematic undercounting of minority viewers. The Coalition, which includes the National Latino Media Council (NLMC), the Hispanic Federation, and 100 Black Men, launched an aggressive campaign to delay the planned roll-out of the LPM's in the New York market. The Coalition's efforts were supported as well by News Corporation, the National Association of Broadcasters, Congresswoman Hilda Solis, Senator Hillary Rodham Clinton as well as many other Congressional representatives, and 20 members of the New York City Council. The Coalition's advocacy paid off and led to the delay of the LPM roll-out in the New York market until June 3, 2004. Meanwhile, the Coalition continues its efforts to delay the roll-out of the LPM in the Los Angeles and Chicago markets.

Fourthly, Nielsen received two stunning blows from the Media Ratings Council (MRC), an independent agency authorized by Congress to audit media ratings services like Nielsen Media Research. The first blow came from the MRC's declaration that Nielsen's proposed use of a household language measure to weight Latino television ratings was flawed. As described by the following news story in Television Week: [2]

> *"…a new debate has been sparked by the release of a white paper by the prestigious Media Ratings Council. The report concluded that the way Nielsen plans to weight its ratings – by households as opposed to the characteristics of individuals residing in those households – is flawed, and not just for Hispanics but for any of the demographic characteristics for which it has already begun weighting."*

The MRC also declared recently that it would withhold accreditation for Nielsen Media Research's local people meter service in New York.[3] Although the lack of accreditation may not delay the roll-out of the LPM's in New York, the failure to obtain accreditation was nonetheless perceived as a blow to Nielsen's plans.

Lastly, other key industry networks have more recently voiced their concerns or support for the local people meter. Univision Communications, the only broadcast network that has not signed with Nielsen Media Research for local people meters, reported that they would continue to hold out in using the LPM results until Nielsen improved the composition of Latino households included in their sample.[4] Support for the LPM was voiced by Viacom's Black Entertainment Network (BET).[5] Although Nielsen claimed support from the NAACP as well, a representative of the NAACP subsequently denied such support. Nielsen spokesman Jack Loftus later apologized for the misunderstanding.

Nielsen Responds Defensively

In response to the mounting criticism, Nielsen Media Research assumed a defensive posture by attempting to discredit its critics and dismissing the possibility of any flaws in their methodology. As Nielsen spokesperson Jack Loftus was fond of saying to Nielsen's critics -- "your baby is ugly" -- suggesting that the LPMs were accurately reporting declines in ratings for television programs that were previously over-rated by the diary method.

Nielsen Media Research also wasted little time in posting a critique of the Latino Television Study on their web site (www.nielsenmediaresearch.com). Table 1 below

summarizes the criticisms made by Nielsen along with the actual facts as described in the study methodology. Had Nielsen staff bothered to read the study methodology, none of these criticisms may have been necessary. Indeed, they were all false and easily verifiable by reading a copy of the full study that was posted on several of the NLMC member web sites, including the National Hispanic Media Coalition web site (www.nhmc.org) when the study was first released.

Table 1
Nielsen Critique of Latino Television Study

Nielsen Criticism	Methodology Reality Check
Study claims to represent entire U.S. Latinos	Generalization limited to Latinos in four markets (Los Angeles, New York, Miami, and San Antonio)
Study relied on listed telephone numbers	Study used both listed and RDD numbers
Fewer than 1 in 9 Latinos (11%) participated in first wave of survey	Actual response rate was 28%, typical of telephone surveys
Latinos who watched a lot of TV were over-sampled, those watching little TV were under-sampled	Latinos were included regardless of their TV viewing frequency
Rincon used unreliable research methods to collect information	Telephone survey method is very reliable and used by national pollsters
Study started with built-in bias that Nielsen TV ratings undercount Latino audiences	The potential undercount was a hypothesis, not a bias

As explained by the Methodology Reality Check, the study findings were generalized only to Latinos in the four target markets, not all U.S. Latinos. Both listed and random digit dialed (RDD) telephone numbers were utilized to contact all selected households to ensure the inclusion of Latino households with unlisted telephones as well as those without Spanish surnames. The actual response rate realized in the study was 28 percent, which is quite typical for studies using the telephone method. Latinos were included in the study if they self-identified as a recognizable Latino group (using Census Bureau categories) and were 16 years or older – their television viewing frequency was not a factor in their selection. The notion that the telephone survey method is unreliable was surprising given that the method is used frequently by national pollsters to gauge the sentiments of U.S. consumers. Indeed, even Nielsen Media Research utilizes the telephone method to conduct its telephone coincidental studies as well as to develop its

home language universe estimates for Latino households. Lastly, what Nielsen defines as a built-in bias actually reflects the hypothesis that the study was exploring.

Two additional points should be made about the accuracy of the telephone method for estimating television audiences. Telephone coincidental research confirms a compliance rate exceeding 90 percent between what a person reports they are viewing by telephone with what the people meter reports that they are actually viewing on television.[6] Also, in a study that compared five different methods for surveying persons in a large metropolitan area, the telephone survey method yielded the closest overall match to the population as determined by comparing the demographic characteristics of the un-weighted sample to similar Census Bureau characteristics.[7] Moreover, the use of a random sample was decidedly more important than the survey response rates in predicting data quality. Consequently, one can have confidence that the telephone survey method reliably measures television viewing or other consumer behavior.

Lastly, in what appeared to be a test of the *Mexican crab theory*, Nielsen Media Research hired the Tomas Rivera Policy Institute, a well-known Latino think tank, to audit the Latino Television Study. By pitting two Latino organizations against each other, Nielsen apparently thought that national attention would be deflected away from their own crisis in methodology. Although the status of this audit remains unclear, it did little to address concerns about Nielsen's methodology and has generated a good deal of ill will in the Latino community.

As if they needed any more bad news, the defensive posture assumed by Nielsen Media Research motivated several members of Congress to request an audit of the Nielsen methodology by the Government Accounting Office (GAO). The GAO audit would focus on comparing the results of the LPM and diary method in selected markets to determine which method is more accurate in estimating television audiences. However, the GAO has since decided that they cannot audit Nielsen Media Research because it is a monopoly, which led to the matter being referred by the Chairman of the Telecommunications Committee to the Judiciary Committee. However, regardless of which method of measurement proves to be more accurate, it would be extremely important for the audit to also focus on the composition of the sample in each market to ensure the inclusion of a good mix of respondents by such demographic attributes as race/ethnicity, gender, nativity and age. Inaccuracies stemming from sampling issues may indeed overshadow the accuracies attributed to the method of measurement.

In the meantime, the Don't Count Us Out Coalition continues its efforts to delay the roll-out of the local people meters in the New York, Los Angeles and Chicago markets. Nielsen Media Research, on the other hand, appears intent on holding their ground.

Nielsen is Not Alone

Although much of the recent national attention has focused on concerns about the Nielsen methodology, it is certainly not the only research organization that could benefit from an external audit. Indeed, other organizations that conduct studies of Latinos with high-stakes consequences include Scarborough, Simmons Market Research, Arbitron and others. Even published studies in reputable journals may require a second look for potential *methodological biases*.

For example, one study of 648 Latinos published in the *Journal of Marketing Research* and often quoted by the Spanish-language media concluded that Spanish-language commercials were more persuasive than English-language commercials.[8] However, the study investigators did not seem concerned about the biases introduced by including respondents who were primarily foreign-born, nor by the elimination of Latinos who spoke only English. Whether by design or oversight, such sampling biases are common in Latino-targeted studies and literally "stack the deck" in favor of Spanish-language media options. Is it any great surprise that Spanish-language commercials were found to be more persuasive than the English-language commercials?

The industry's constant obsession with *affirming the virtues of Spanish-language media* presents another barrier that limits the usefulness of Latino-targeted studies. Such studies continue to affirm that Spanish-language media are more important to Latinos, contribute to their political and economic development, and are more effective in communicating the appropriate cultural values to Latino families. By contrast, English-language media are usually evaluated more negatively due to the use of negative Latino stereotypes, the exclusion of Latinos, and the proliferation of unsuitable content. This obsession with Spanish-language media is further reinforced by the industry's spending practices – 95 percent of total ad billings in 2002 were spent by marketers on Spanish-language media[9] -- a practice that tacitly ignores the existence of native-born Latinos who primarily depend on English-language media. Interestingly, the Federal Trade Commission recently created a Hispanic Law Enforcement and Outreach initiative to monitor the significant growth of fraudulent and unethical advertising in Spanish-

language media.[10] Perhaps the industry's obsession in extolling the virtues of Spanish-language media may be better invested in Latino studies that are more scientifically sound.

Eventually, such biased studies tend to *mislead* advertisers who may wonder why their campaigns do not hit their intended mark. For example, various news stories tell us about the aggressive outreach efforts launched by both political parties to reach the coveted Latino vote and the total reliance on Spanish-language advertising to reach that objective. What the political campaigners may not realize, however, is that the majority of Latino voters are native-born residents who rely primarily on English-language media for their information needs.[11,12] Perhaps the campaigns will discover this concept before the end of the election season.

Latinos Are Difficult to Research

Undoubtedly, the challenges of researching U.S. Latinos are numerous. Companies like Nielsen Media Research that require stable measurements over time encounter major obstacles in large urban markets with *large renter populations* like Los Angeles (63% renters) and New York (78% renters).

Latinos, especially recent immigrants, tend to be *distrustful of U.S. institutions* and require time to overcome their fears. Indeed, the prospects of having their homes monitored electronically are not easily overcome, even with current monetary incentives.

The *varying language abilities* in many Latino households creates a tremendous staffing responsibility for the research company, including bilingual supervisors, coders, recruiters, interviewers, and analysts. Moreover, the lower educational level of Latinos in many urban markets points to potential problems with written tasks, whether in English or Spanish, and the corresponding need to lower the reading difficulty level of instructional materials.

The ability of the researcher to *select a representative sample of Latinos* is further challenged by their residential patterns. Foreign-born Latinos, for example, tend to cluster in high-density zip codes while native-born Latinos are more dispersed geographically. An inappropriate sampling strategy could yield a disproportionate number of foreign-born or native-born Latinos and thus bias the study findings.

Common Sources of Bias in Latino Studies

Given the challenges discussed above, Latino-targeted studies are most likely to suffer from one or more of the following problems: sample selection bias, unreliable measurement instruments, monolingual interviewers, and improper sample weighting.

Biased Sample Selection*:* Biased sample selection is among the worst violators and occurs in several ways. When a sample is selected by focusing exclusively on geographic areas with *high concentrations of Latinos*, it biases the sample towards lower income, Spanish-dominant immigrants who depend primarily on Spanish-language media. Another common practice is to select the sample from the universe of *Spanish-surnamed households*, a procedure that can provide uneven results. For example, our experience suggests that a Spanish-surname is about 85 percent accurate in identifying a Latino-origin person while the other 15 percent are excluded for lack of a commonly recognized Spanish surname. The quality of the Spanish-surname list used by sampling shops varies a great deal. Some of these lists may include a broad list of Spanish surnames, while others may be limited to the most commonly recognized names in a particular geographic region. A sample may under-represent Latinos in some regions simply because the sampling shop used an outdated or incomplete list of surnames. The increasing intermarriage rate between Latinos and non-Latinos means that an increasing number of Latinas will be excluded by relying exclusively on Spanish surnames. To overcome these shortcomings, researchers have typically relied on RDD (random digit dialed) samples, which are computer-generated telephone numbers that theoretically include all listed and un-listed telephone numbers in a designated area. When used in Latino-targeted studies, however, RDD samples are highly inefficient and costly since they require the dialing of thousands of non-working banks of telephone numbers. In some cases, studies using these two sampling techniques have produced similar results. To address these shortcomings, we have used a sampling technique that combines the listed surnames and RDD technique to produce a more representative sample of Latinos. Regardless of the sampling technique utilized, the bottom line should focus on the quality of the un-weighted sample realized – it should match known Census Bureau demographics as closely as possible for the targeted consumer and geographic area.

Yet another selection bias involves the *pre-screening of Latinos by language ability or media consumption habits*. For example, a study of Hispanic Internet usage by the Pew Internet Project reported in press releases an Internet usage rate of 50 percent,

a rate that was twice as high as the rate reported by the Census Bureau in 2000 (24%).[13] What was not immediately obvious in the press releases and ensuing stories was that the study included only English-speaking Latinos. Because the Internet usage rate among English-dominant Latinos is known to be decidedly higher than Spanish-dominant Latinos, the public was misled into thinking that all Latinos were embracing the Internet at a higher rate that previously thought.[11] A similar problem occurred in one national study of Latinos that was sponsored by a Spanish-language television network. Latino respondents were pre-screened on the basis of having viewed five or more hours of Spanish-language television and were subsequently queried about their usage of Spanish-language media. The conclusion that Latinos viewed a lot of Spanish-language television was not totally unexpected given the selection bias, but the investigators proceeded to generalize the study results to <u>all</u> U.S. Latinos – which was incorrect given their sampling procedure.

<u>Unreliable Measurement Instruments</u>: Our industry is more familiar with the biases that can result from *surveys that include poor translations, vague wording, and higher reading difficulty levels* than necessary. However, because 60 to 70 percent of U.S. Latinos have not graduated from high school, these biases can present significant barriers whether the surveys are presented in *English or Spanish*.

In the case of television viewing, recent immigrants are especially impacted by the requirement to complete *written diaries* to record their viewing behavior. Our own experience confirms that immigrants have better recall of Spanish than English names, whether it involves television programs, street names or cities. In the case of written diaries, the dependence on retrospective recall of television programs thus poses an additional burden for Latino immigrants, perhaps leading to lower reported viewing of English-language television programming.

Local people meters, wherein a household member is required to press a button to indicate that they are viewing or not viewing a television program, are generally considered more accurate than written diaries for recording viewing behavior. Until recently, the prohibitive costs of local people meters had limited their distribution in all Nielsen households, although the New York market is slated to begin using people meters on June 3rd, followed by the Los Angeles and Chicago markets. As discussed earlier, however, the planned roll-out of the local people meter has met stiff resistance from community organizations and selected members of Congress. Although Nielsen Media Research maintains that people meters accurately measure Latino and African-

American television audiences, they have not identified any studies to support this claim. The notion that local people meters play a role in lower television ratings is not without merit. For example, one study revealed that fatigue and conditioning could lead to lower audience ratings, although no comparisons were provided by race or ethnic group.[14] In her review of the people meter wars, Professor Karen Buzzard from the Media, Journalism, and Film Department of Southwest Missouri State University, summed up her analysis of the drop in ratings and use of people meters:

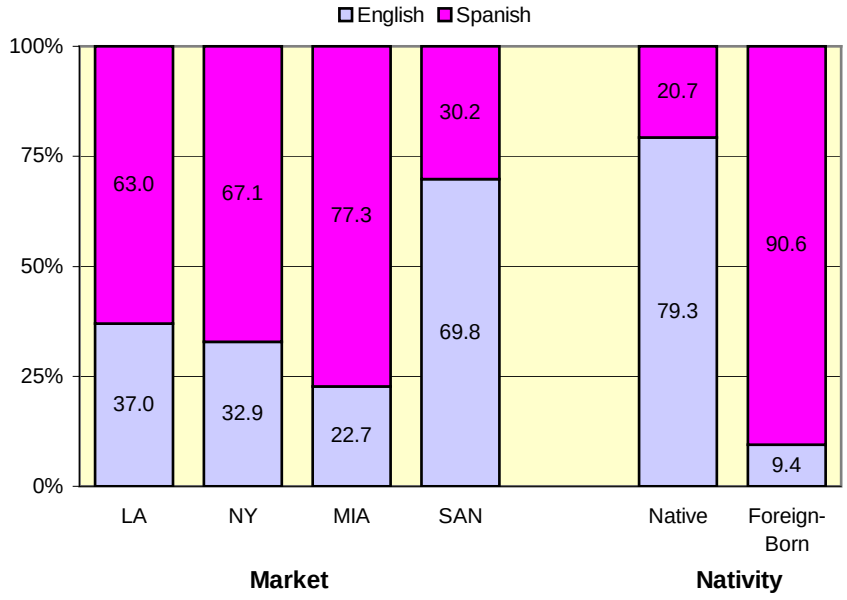
> "Overall, People meters have resulted in lower viewing levels because they re-introduce the need for viewers to actively participate in the ratings process and sharply lower kids and teens ratings – the audience segments least likely to participate." [15]

Given these concerns about the people meter as well as the fact that Latino households include a higher proportion of youth than non-Latino households, it would seem like a good idea for Nielsen Media Research to conduct a study that rules out the possibility that the lower television ratings result from differences in the manner that the people meter is used by Latinos and African-Americans.

Monolingual Interviewers: The use of monolingual interviewers can undermine even the best study designs. Monolingual interviewers tend to dictate the type of respondent selected in a Latino study by selecting respondents that more closely match their own language skills. For example, Spanish-monolingual interviewers are more likely to interview Spanish-dominant respondents, while English-monolingual interviewers tend to interview English-dominant Latinos. Only bilingual interviews can ensure that both English-speaking and Spanish-speaking Latinos will have the same opportunity of being selected for the interview. The problem occurs when the initial contact is made to the household. A Latino respondent who hears an interviewer greeting in Spanish may not be prepared to respond in Spanish, and will likely seek out another household member who is proficient in Spanish or just refuse to conduct the interview. A similar dynamic is noted when a respondent hears an English-monolingual interviewer – if not proficient in English, the respondent will either refuse to cooperate or seek out another household member who is proficient in English. The respondent has no way of knowing that the interview can be conducted in either English or Spanish unless explicitly told by a bilingual interviewer. Even after this option is made explicit to the respondent, it is good practice to ask the Latino respondent if they would prefer to

continue the interview in English or Spanish. The language that Latinos choose when given the choice represents, in our view, their dominant language and is more likely to produce reliable and valid interviews. As a case in point, Figure 1 below presents the language choices made by Latinos when provided the choice by a bilingual interviewer, as reported in the Latino Television Study.

Figure 1: Interviewing Language by Market and Nativity



Source: Latino Television Study, National Hispanic Media Council, 2004.

The chart clearly shows that eight in ten (79.3%) of the native-born Latinos chose English as the interviewing language when given the choice, while over nine in ten (90.6%) of the foreign-born Latinos chose Spanish. The choice of English as the interviewing language was notably higher in the San Antonio market (69.8%), while the preference for Spanish was greater in Los Angeles (63.0%), New York (67.1%), and Miami (77.3%). It is clear that a monolingual interviewer would have materially changed these language choices as well as the study outcomes. A well-trained bilingual interviewer can greatly minimize this type of interviewing bias, although one important caveat is needed here. Interviewers need to be properly screened for their reading and writing skills in English and Spanish and excluded from a Latino study if they do not reveal a balanced set of such skills in both languages. Otherwise, interviewers who show weak linguistic skills in one language will tend to exclude Latinos who prefer to be

interviewed in that particular language. On-going monitoring of all interviewing staff throughout the life of a study is the only way to ensure that the interviewers are not inadvertently biasing the study outcomes. Because a good bilingual interviewing team is critical to the successful completion of a Latino-targeted study, their compensation should be commensurate with the skills and responsibilities inherent in such studies.

Varying Cooperation Rates: Despite one's best efforts, many Latino studies end up with higher proportions of female and foreign-born Latinos than would be expected from Census demographics. Why? Because females are more cooperative than males, and foreign-born Latinos are more cooperative than native-born Latinos. The subjective biases of interviewers can easily lead to such imbalances as well when supervisors are lax in monitoring the interviewing staff. By contrast, significant cooperation problems are encountered among Latino households, particularly among the foreign-born, by ratings services like Nielsen that depend on household panels. Moreover, the higher turnover rate among Latino households continues to threaten the integrity of the sample and accelerates the cost of research. Significant sample imbalances can result in Latino studies that overlook the problems caused by low cooperation rates.

Weighting Problems: When the sample characteristics of a survey vary noticeably from known characteristics, i.e., Census Bureau demographics, it is customary to use post-stratification weights to correct for these sample imbalances. Weighting problems in Latino studies are more likely to occur in two ways: *when weighting is not used when needed*, *or unreliable/incorrect weights are used*. To illustrate the first type of problem, Table 2 on the following page un-weighted and weighted ratings that resulted from the Dallas/Ft. Worth Latino Trendline study of 600 Latino adults who were queried about the television network that they watched most often for news.[11]

Table 2
Un-weighted vs. Weighted Television Ratings

(Percents)

	Network Watched Most Often for News				
	Univision	*Telemundo*	*Fox*	*ABC*	*NBC*
Un-weighted Percent	61.1	10.0	7.9	6.4	5.7
Weighted Percent (Based on gender, nativity in Census 2000)	52.0	9.0	11.7	8.3	7.5
Rating Difference	-9.1	-1.0	+3.8	+1.9	+1.8

Source: Dallas/Ft. Worth Latino Trendline, Rincon &Associates, 2004.

One can readily see from the row labeled "Un-weighted Percent" in Table 2 above that 61.1 percent of the Latino adults viewed Univision most often for news. However, the percentage of Latino adults viewing Univision drops to 52.0 percent after applying the weights to the viewing data – a drop of –9.1 percent that translates to a loss of 87,360 viewers. By contrast, the "Weighted Percent" for each of the English-language networks gained audience shares as one would expect in a sample that under-represents native-born Latinos. If the Univision rating discrepancy had been 15 percent instead of 9.1 percent, the potential loss of viewers could approach 144,000. At the national level, these differences can translate to thousands of Latino viewers for any given network. Thus, the failure to use weighting when it is needed can lead to dramatically different audience shares for the media alternatives under consideration.

It is also likely for Latino studies to use *unreliable or incorrect weights*. The household language classification scheme used by Nielsen Media Research represents a good example of unreliable weighting. Table 3 on the following page presents the language categories and related classification rules used by Nielsen Media Research to weight their Latino television ratings.

Table 3
Nielsen Home Language Classification Scheme

Language Group	*Classification Rule*
Only Spanish	All household members report that they only speak Spanish in the home
Mostly Spanish	These are homes that report having a mixture of people present who speak Only Spanish, Mostly Spanish or **Spanish and English Equally**. Each language group is counted as present if it has at least one representative in the home
Only English	All household members report that they only speak English at home
Mostly English	These are homes that report having a mixture of people present who speak Only English, Mostly English, or **Spanish and English Equally**. Each language group is counted as present if it has at least one representative in the home
Spanish/English Equally	**These are all homes that do not fit into the above categories**. For instance, a home that reported having at least one Mostly English and one Mostly Spanish Speaking home would be considered a Spanish/English Equal home

Source: Doug Darfield, VP Hispanic Services, Nielsen Media Research, 12-05-03.

The classification rule utilizes the home language spoken by each household member (2 years or older) and sorts the household into one of the five language groups. If, for example, the proportion of Latino households in the Spanish Only category falls short during any measurement period, then Nielsen re-aligns or weights the proportion of Spanish Only households according to what they think it should be – not the Census Bureau, but an independent universe language estimate that Nielsen Media Research has created from several thousand telephone interviews of U.S. Latino households.

There are two problems with using the household language categories as a weighting variable for Latino television ratings. First, the categories themselves are vague and redundant, especially the Mostly Spanish and Mostly English categories where both include persons who speak Spanish and English equally. The Spanish/English Equally category sounds more like a default bin to place household members that are difficult to place anywhere else. The classification scheme is analogous to taking the blood type for every household member and trying to come up with one blood type for the household – just because you can do it does not make the results meaningful. The second problem relates to the use of the household language weight when the ratings are based on the number of people. As mentioned earlier in this

report, the Media Rating Council had also concluded that Nielsen's household weighting system was flawed for any of the demographic characteristics for which it had already begun weighting.

The Latino Television Study: A Solid Study Design

Recognizing the great potential for under-counting television audiences, the National Latino Media Council (NLMC) embarked on a mission to get "a second opinion" by initiating the Latino Television Study during August of 2003. In a nutshell, the study was designed to address four key areas:

- Assess the reliability of the Nielsen household language measure
- Describe Latino television viewing patterns
- Evaluate potential discrepancies in the television viewing audiences derived by Nielsen and the present study
- Describe the future composition of Latino television audiences

The reader is encouraged to read the full study, including a detailed description of the study methodology, on the National Hispanic Media Coalition (NHMC) web site (www.nhmc.org). In the discussion that follows, we will review aspects of the study design that enhances its methodological soundness and present several key study findings.

Study Design: The Latino Television Study was completed in three waves from August to October of 2003 in four ADI markets: Los Angeles, New York, Miami, and San Antonio. The first wave included telephone interviews with 1,536 Latinos who were queried on a broad range of media usage, home language usage, and demographic attributes. The second wave of 311 Latinos was conducted with a second household member within one week after the first wave interviews were completed in each market. Second wave telephone interviews were conducted to determine the extent to which the description of the home language usage for different household members agreed with the descriptions reported earlier by the first-wave respondents. Lastly, 195 third-wave interviews were completed 7-10 days after the second-wave interviews with either the first-wave or second-wave respondent to evaluate temporal changes in the descriptions of the home language used by different household members. Following is a list of specific study design characteristics:

- The study design, data collection, data analysis and written report were completed by Rincon & Associates
- Both listed surname and RDD techniques were utilized to select a simple random sample in each market
- A total of 384 telephone interviews were completed in each market with a margin of error of plus or minus 5 percent at a confidence level of 95 percent
- Respondents were screened on age (16 years or older) and self-identification as a Hispanic or Latino using Census Bureau categories
- The recent birthday technique was used to randomly select a respondent within each household
- A pilot test was conducted prior to data collection activities
- Only bilingual interviewers were employed in the study and monitored closely by bilingual supervisors
- Survey respondents were provided the choice of the interviewing language, regardless of which language they used to answer the telephone
- The response rate of 28 percent realized by the first wave of the study was typical of telephone-based surveys
- The study data was weighted by gender and nativity using the Census Bureau 2000
- Prior to the study's release, survey research and sampling experts reviewed the final report to identify needed modifications

These study design characteristics confirm the fact that great care was taken to avoid the common sources of bias associated with Latino studies, thus providing the reader added confidence in the study findings.

Key Study Findings: Following are some of the key findings from the study, although the reader is encouraged to review the full report on the NHMC web site.

The home language measure used by Nielsen Media Research to classify Latino households into language categories was unstable when used by different household members. For example, the agreement rate (see Table 4 on following page) was 57.1 percent between two different household members (308 pairs) in describing the home language usage of Person 2 in their household. In general, the agreement rate decreases with additional household members, pointing to an unstable measure. The lack of stability is evident regardless of whether a full or collapse language scale is used.

Table 4
Agreement Rates on Home Language Measure Between Two Household Members

Household Member	Number of Comparisons	Percent Agreement	
		Full Scale	Collapsed Scale
Person 1	311	60.8	82.6
Person 2	308	57.1	79.9
Person 3	221	53.8	67.4
Person 4	144	49.3	68.8
Person 5	80	52.5	70.0
Person 6	38	47.4	60.5
Person 7	13	30.8	53.8
Person 8	4	50.0	50.0
Person 9	3	100.0	100.0
Person 10	1	50.0	100.0

Source: Latino Television Study, National Hispanic Media Council, 2004.

The Nielsen home language measure has one very problematic category. A discriminant analysis was conducted to address the question – If you know how well a person understands English and Spanish, how much they use English and Spanish-language media, and their gender, age, years of U.S. residency, education and nativity – how accurately could you classify that person into one of the Nielsen language categories? Table 5 below presents the classification results of the discriminant analysis.

Table 5
Classification Rate for Nielsen Language Groups

Nielsen Language Group	*Correct Classification Rate*
Spanish Dominant	78.3%
English Dominant	77.3%
Spanish/English Equally	32.1%

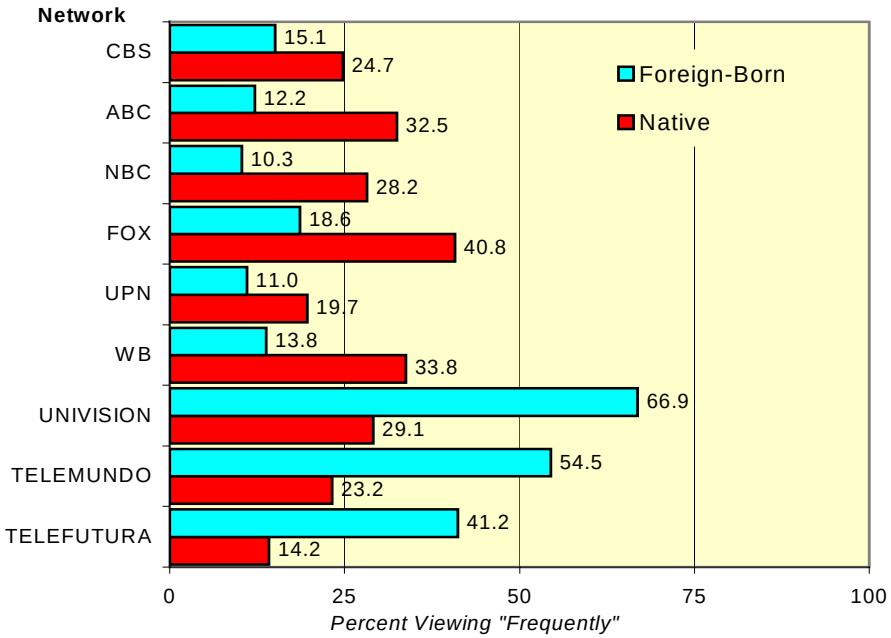
Source: Latino Television Study, National Hispanic Media Council, 2004.

The correct classification rates show a fairly high hit rate for the Spanish Dominant (78.3%) and English Dominant (77.3%), but a very poor hit rate of 32.1 percent for the Spanish/English Equally group. This suggests that the respondents in the Spanish/English Equally group share less in common than the other two groups in terms of the attributes included in the analysis. This finding is not surprising given the vagueness and redundancy of the categories discussed earlier.

More suitable alternatives for weighting television ratings are available in the decennial census by the U.S. Census Bureau. A series of regression analyses were conducted to address the question: If weighting must be used, what else could be substituted for the unstable home language measure? Average weekly hours dedicated to English-language television and Spanish-language television were used as the dependent variables in the regressions, while the independent variables included nativity, home language spoken, ability to understand English, and years living in the U.S. The results of the regressions indicated that nativity and ability to understand English were very useful in predicting the average amount of time spent viewing English and Spanish-language television, and would serve as better measures to adjust Latino television ratings than Nielsen's home language measure. Moreover, these two variables are available from the Census Bureau for all geographic areas in the U.S.

The core viewing audience for English-language television networks were native-born Latinos, while the core television audience for Spanish-language networks were foreign-born Latinos. As illustrates by Figure 2 below, although cross-viewing of television networks was evident for both native-born and foreign-born Latinos, Spanish-language networks primarily appealed to foreign-born Latinos while the English-language networks primarily drew a native-born Latino audience.

Figure 2: Frequent Viewers of Network Television by Nativity



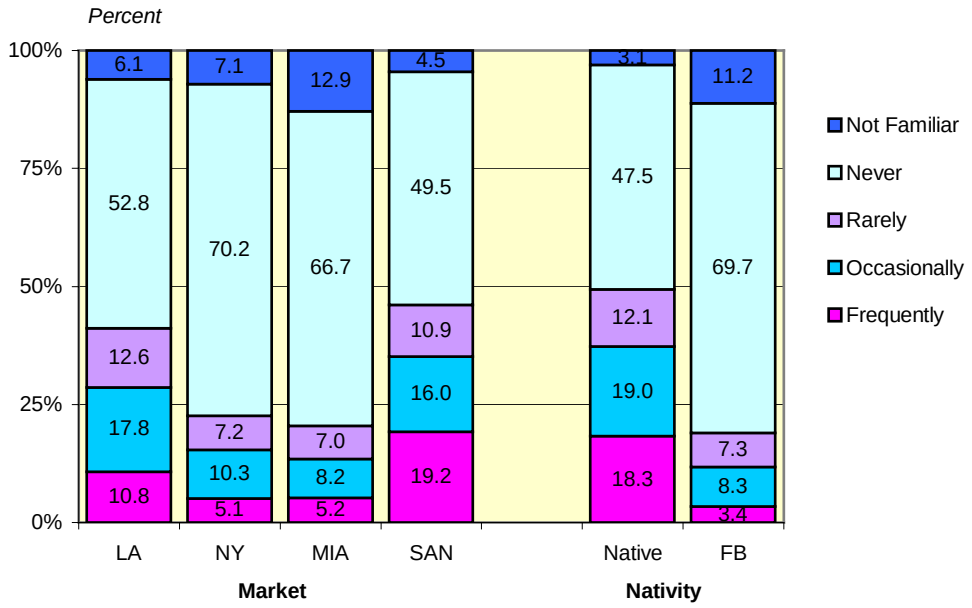
Percents based on multiple responses

Source: Latino Television Study, National Hispanic Media Council, 2004.

The advantages enjoyed by the Spanish-language networks over the English-language networks included a strong Latino appetite for novelas, the strong Latino dependence on the Spanish language, and the paucity of theaters and video rental stores that can meet their Spanish-language entertainment needs.

The core television audience for The George Lopez Show consisted primarily of native-born Mexicans. As shown by Figure 3 below, native-born Latinos were over three times more likely (37.3%) than foreign-born Latinos (11.7%) to watch The George Lopez Show either frequently or occasionally. The show was also considerably more popular in markets with higher concentrations of Mexican-origin Latinos, such as Los Angeles and San Antonio.

Figure 3: Frequency of Viewing The George Lopez Show by Market and Nativity



Source: Latino Television Study, National Hispanic Media Council, 2004.

The Latino audience for English-language television programs may be significantly under-estimated by Nielsen Media Research. Table 6 below presents our audience estimates for The George Lopez Show. Based on an estimated viewing audience of 8.9 million Latinos who were 16 years or older, approximately 802,553 Latinos watched the show frequently, while an additional 1,257,723 Latinos watched the show occasionally -- which totals to a viewing audience of two million Latinos. This compares to the Nielsen estimated Latino audience of 1.2 million who watched The George Lopez Show in 17 Latino markets – a glaring discrepancy that begs for further disclosure of Nielsen's audience estimation procedures.

Table 6
Audience Estimates for The George Lopez Show

(Latinos 16 years or older)

Viewing Frequency	Market				Total Viewers
	Los Angeles	*New York*	*Miami*	*San Antonio*	
Frequently	464,891	135,698	66,102	135,862	802,553
Occasionally	766,209	274,057	104,238	113,219	1,257,723
Total Viewing Frequently/ Occasionally	1,231,100	409,755	170,340	249,081	2,060,276

Source: Latino Television Study, National Hispanic Media Council, 2004.

The George Lopez Show appeals primarily to second and third-generation Latinos. As Table 7 below reveals, the popularity of The George Lopez Show is very evident among second and third-generation Latinos, but considerably lower among first-generation (foreign-born) Latinos. Thus, the size of the Latino audiences for shows like The George Lopez Show will depend greatly on the extent to which Nielsen includes second and third-generation Latinos in their research panels.

Table 7
Differential Appeal of The George Lopez Show by Generational Status

(Latinos 16 years or older)

Viewing Frequency	**Generation**		
	First	*Second*	Third
Frequently	3.4%	18.5%	18.4%
Occasionally	8.3%	15.0%	22.8%
Total Viewing Frequently/ Occasionally	11.7%	33.5%	41.2%

Source: Latino Television Study, National Hispanic Media Council, 2004.

The rise of the second generation will profoundly impact the composition of future Latino audiences and the demand for English-language programming in the film and television industries. In their paper entitled "The Rise of the Second Generation: Changing Patterns in Hispanic Population Growth," Roberto Suro and demographer Jeffrey S. Passel explained the shifting composition of the Latino population and pointed out that the majority of Latino growth between 2000 and 2020 will come from second and third generation Latinos – the children of immigrants.[16] Table 8 below shows that Latino growth during the 1970 to 2000 period was fueled primarily by first-generation Latinos (45%), although this segment is projected to comprise only 25 percent of the estimated Latino growth between 2000 and 2020.

Table 8
Generational Shift in Latino Population Growth

	1970-2000	2000-2020
Latino Population Growth	9.6 million to 35.3 million	35.3 million to 60.4 million
Net Increase	25.7 million	25.1 million
First Generation	45%	25%
Second Generation	28%	47%
Third Generation	27%	28%
	100%	100%

Source: Pew Hispanic Center, 2004

As the more English-dependent children of immigrants displace the foreign-born as the primary driver of Latino growth into the future, the demand for shows like The George Lopez Show that appeal to second and third-generation Latinos will increase dramatically.

Other Study Findings

- Latinos were more likely to view television without the presence of children. Over half of Latino children primarily viewed English-language television networks.
- Latinos had broad access to both English-language and Spanish-language television programming. Half of all Latinos had access to cable television, while less than two in ten Latinos had access to satellite television. Simultaneous audio programming (SAP) was used by two in ten Latinos, especially the foreign born.
- In 2003, U.S. Latinos spent an estimated $12 million in the entertainment industry and $680 million on movie admissions, representing a significant economic investment.
- Regardless of nativity, the movies most recently viewed by Latinos included a variety of English-language genres such as horror, comic heroes, comedies and action movies.
- Native-born Latinos were less likely than foreign-born Latinos to enjoy a movie that included dubbing in Spanish, Spanish subtitles, and non-Latino actors playing Latino roles.

Suggested Focus for Nielsen Audit

The study pointed to a potentially large under-count in the Latino viewing audience for The George Lopez Show when compared to the audience reported last fall by Nielsen Media Research. According to Nielsen Media Research, the estimated Latino viewing audience for The George Lopez Show was 1.2 million viewers in the 17 markets analyzed during the period Sept. 9-22, 2003. By contrast, the Latino Television Study estimated 802,553 Latinos who viewed the show "frequently" and 1,257,723 who viewed the show "occasionally" (Aug.-Sept. 2003). Most importantly, the Latino Television Study included only four of the 17 Latino markets analyzed by Nielsen Media Research.

The observed discrepancy between the two estimates is substantial and requires a thorough review of the procedures used to derive both estimates. The size of the discrepancy cannot be explained alone by the two different methods of measurement employed (telephone vs. diary method) since the compliance rates between these two methods is known to exceed 90 percent. Since the Latino Television Study estimation procedure has already been disclosed, it is equally important that Nielsen Media Research also disclose the details related to their audience estimation procedures.

An audit by an independent body should compare the accuracy of the local people meter with the traditional diary method. Regardless of which method proves to be more accurate, however, the discrepancy related to The George Lopez Show cannot be resolved unless Nielsen Media Research is required to address the following questions for the measurement period of Sept. 9-22, 2003 and all 17 Latino markets included in this period.

Sampling Issues

- What was the composition of the <u>un-weighted sample</u> of Latinos in terms of gender, age, and nativity?
- How was the sample distributed geographically?
- How were the Hispanic households selected?
- What was the cooperation rate for Latino households? For non-Latino households?
- What was the retention rate for Latino households? How do these rates compare to non-Latino households?

Data Collection Issues

a. Diary Method

- What instructions are provided to Latino families to complete the diaries?
- Are bilingual interviewers utilized in all recruitment and interviewing activities? What proportion of the interviewers are foreign vs. native born? How are their language English and Spanish-language skills evaluated?
- What instructions are provided to family members in regards to recording the viewing behavior of children 2 to 12 years old?
- What are the reported compliance rates for the written diaries? What are the compliance rates for Latinos vs. other racial/ethnic groups?
- Have any telephone coincidental studies been completed to verify compliance rates for Latino and non-Latino households?

b. People Meter

- Studies show that fatigue and conditioning can lead to a decline in the audience ratings derived from people meters. Have you conducted such a study?
- Have any studies been conducted regarding the differential effects of fatigue and conditioning on the ratings provided by Latinos and non-Latinos when using people meters?
- What are the compliance rates for the people meter by age? Are teenagers and younger children less likely to use the people meter reliably than older members of the household? Do Latino teens and younger children use the people meter as reliably as non-Latino teens and children?

Weighting

- What weighting formulas were used to transform the raw data into the final ratings for The George Lopez Show during the period of Sept. 9-22, 2003? Please explain the procedure, variables and the population sources included in the weighting formulas.

Other Questions

- In a New York Times article (4-5-04), Nielsen Media Research reportedly admitted to under-counting 300,000 Latinos in the New York market during 2000. How did this under-count occur? How was it discovered, and how was it corrected?
- A white paper conducted for the Media Rating Council reported that the household language weighting procedure proposed by Nielsen Media

Research was flawed. Please explain why they reached this conclusion, as well as provide a copy of the white paper.

Restoring Confidence in the Nielsen Methodology

A recent editorial in The New York Times entitled "Who Tests Voting Machines?" revealed striking similarities to the calls for accountability in the broadcast industry.[17] For example, when questions are raised about the reliability of electronic voting machines, election officials reassure the public that independent testing authorities ensure that the machines are accurate and honest. The "independent" testing companies routinely deny government officials and the public basic information about their testing procedures or testers' credentials. Interestingly, while they are called independent, the testing labs are selected and paid by the voting machine companies, which exert considerable pressure to review things quickly and not to find problems. To radically improve the voting system, the author recommends the following:

- **Truly independent testing** laboratories that are paid by the government, not the voting machine companies
- **Transparency**, which explains to voters how the testing is being done and the tester's qualifications
- **Rigorous standards** that spell out in detail how software and hardware are to be tested and deficiencies fixed
- **Tough penalties** for voting machine companies and election officials who try to pass off uncertified software and hardware as certified

Each of these recommendations bears some relevance to the Nielsen ratings system. The call for independent testing of the Nielsen ratings by an entity that has no relationship to Nielsen Media Research or its parent company is critical. However, the independent agency must be "transparent" and willing to publicly disclose its audit procedures as well as the credentials of its auditors. In the case of studies that include minority audiences, the skills of the auditors involved in the process should go beyond financial auditing practices to include expertise in survey methodology and the media behavior of minority audiences. Although Nielsen Media Research adheres to the rigorous standards defined by the Media Rating Council, both of these organizations would likely benefit by considering the methodological issues identified in the present study. Lastly, tough penalties for methodological irregularities may be necessary in

those circumstances where media ratings services appear unwilling to comply with accepted industry practices or demands for increased accountability from public officials.

Reference Notes

1. Paterson, D. "Airwave alert." New York Post, April 5, 2004.
2. Mandese, J. "Getting a handle on Hispanic households: Complexity of Spanish-language market makes weighting viewer characteristics a challenge for Nielsen," Television Week, March 29, 2004.
3. Bachman, K. "MRC withholds accreditation for New York LPM," MediaWeek.com, May 28, 2004.
4. Bachman, K. "Univision holds out again LPM," MediaWeek.com, May 25, 2004.
5. Bachman, K. "Nielsen picks up LPM support from the NAACP," MediaWeek.com, May 28, 2004.
6. Danaher, Peter J. and Beed. T.W. "A coincidental survey of people meter panelists: Comparing what people say with what they do." Journal of Advertising Research, January/February 1993.
7. Selzer & Company. "Can you trust your data?" Presentation to the American Association of Public Opinion Researchers, May 2003.
8. Roslow, P. and Nicholls, J.A.F. "Targeting the Hispanic market: Comparative persuasion of TV commercials in Spanish and English," Journal of Advertising Research, May/June 1996.
9. Hispanic Business. "Top 25 Billing Survey," December 2002.
10. Londoño, E. "FTC targets fraud aimed at Hispanics." Dallas Morning News, February 28, 2004.
11. Rincon, E. T. Dallas/Ft. Worth Latino Trendline, Rincon & Associates, 2004.
12. Suro, R. and Passel, J. "Changing Channels and Crisscrossing Cultures: A survey of Latinos on the News Media," Pew Hispanic Center, 2004.
13. Pew Internet & American Life. "Hispanics and the Internet," July 25, 2001.
14. Cook, B. "Debugging the people meter." Presentation at the Television Research – International Symposium, Tarrytown, N.Y. October 15-18, 1989.
15. Buzzard, K. "The people meter wars: A case study of technological innovation and diffusion in the ratings industry," Journal of Media Economics, 15(4), 273-291, 2002.
16. Suro, R. and Passel, J. "The rise of the second generation: Changing patterns in Hispanic population growth," Pew Hispanic Center, 2003.
17. New York Times. "Who tests voting machines?," www.nytimes.com, May 30, 2004.