

Transitioning from Self-Reports to Self-Installed Electronic Audience Measurement

Norman Trussell¹, Lorelle Vanno¹, Elizabeth Matthes¹, Justin Bailey¹, and Michael Link²

¹The Nielsen Company, 501 Brooker Creek Blvd., Oldsmar, FL 34677

²The Nielsen Company, 1145 Sanctuary Pkwy., Suite 100, Alpharetta, GA 30004

Abstract

For the past century, self-reports have served as the primary means of collecting information from the public about different types of behavior. Technological innovations have opened new doors for measuring certain behaviors through electronic means. Nielsen has used self-reports recorded in a paper diary for television audience measurement since the 1950s. Yet, as viewing choices have increased and television technology has evolved, respondents increasingly have difficulty accurately and completely recording all viewing information in a paper-based diary. Over the last several years Nielsen began to leverage these newer technologies with expensive and relatively invasive electronic metering devices (traditionally reserved for national ratings) deployed to replace the diary in the largest local television markets. More recently, Nielsen developed the “Mailable Meter”, a smaller self-installed television meter that captures tuning data (what shows were watched and for how long). This technology can potentially collect more complete and accurate television tuning information than the diary, while reducing respondent burden (completion of a much simpler viewing log of who is watching). Methodologies were developed to maximize respondent cooperation and compliance, focusing on three key areas: (1) recruitment techniques to ensure a high level of commitment among participating households; (2) a support structure to provide assistance to respondents throughout the measurement period; and (3) an optimized incentive structure which balances participation gains with cost. In November 2007, a Mailable Meter field test was conducted 406 households in parallel with the self-reported Nielsen Diary Service measurement. Key metrics from this test were analyzed across different demographic groups, including 1) recruitment rates, 2) return rates, and 3) respondents’ experiences and perceived burden via data collected by a follow-up questionnaire after completion of the study. Findings from this effort are compared with those from the self-reported diary and are discussed within the framework of the growing shift from self-reports to electronic behavior measurement. Based on the results of testing so far, the idea of combining self reporting with electronic measurement appears to be a concept with the possibility of both easing respondent burden and increasing the accuracy of the data collected in the context of collection of television ratings.

Key Words: electronic measurement, self-reports, audience measurement, media behavior, methodology, non-response, attitudes & behaviors

1.0 Background

1.1 The need to move from self-reports to electronic measurement

Historically self-reports have served as the primary means of collecting information from the public about different types of behavior. Nielsen has used self-reports recorded in a paper diary for television audience measurement since the 1950s. While this has traditionally been the only option available in small markets, many researchers agree that self-reports do not always lead to accurate information. According to an article in the *American Psychologist*, “. . . research into behavioral reports consistently demonstrated that mundane and frequent behaviors are poorly represented in memory, forcing respondents to rely on estimation strategies” (Schwarz, 1999, p.100). Additionally, as viewing choices have grown exponentially and television technology has advanced respondents increasingly have difficulty accurately and completely recording all viewing information (including who is watching, for how long, what program was being watched, the channel and call letters of the station) all in a paper-based diary.

1.2 - The Mailable Meter – an overview

In recent years, technological innovations have opened new doors for measuring certain behaviors through electronic means. Over the last several years Nielsen began to replace the diary methodology in the largest local television markets with electronic metering devices installed by field representatives. These meters are considered the most accurate and reliable option for measuring households and people using television, but are very expensive, relatively invasive, and difficult to install. Nielsen has a potential alternative replacement for the diary in mid-sized

to smaller markets in development. This product is known as the “Mailable Meter” and is a smaller self-installed device that captures tuning data through audio signatures when placed near a television set. This technology can potentially collect more complete and accurate television tuning information, while reducing respondent burden with the completion of a much simpler viewing log of who is watching.

2.0 Methods and Design

Mailable Meter methodologies are being developed and tested with a goal of maximizing respondent cooperation and compliance, focusing on three key areas: (1) recruitment techniques to ensure a high level of commitment among participating households; (2) a support structure to provide assistance to respondents throughout the measurement period; and (3) an ongoing program of research on incentive strategies which seeks to reach a balance of participation versus cost to reach an optimized incentive structure.

In November 2007, a Mailable Meter field test was conducted in parallel with self-reported diary measurement with 406 households. The recruitment and shipping process is outlined in the flow chart in Figure 1. Three weeks after the end of the study (on December 26th) participating households were sent a follow-up questionnaire in an effort to collect information about their experiences with the study and thoughts on the process. The follow-up questionnaire was mailed to 391 of the participating households (excluding only 15 households that had opted out or otherwise been removed from the study).

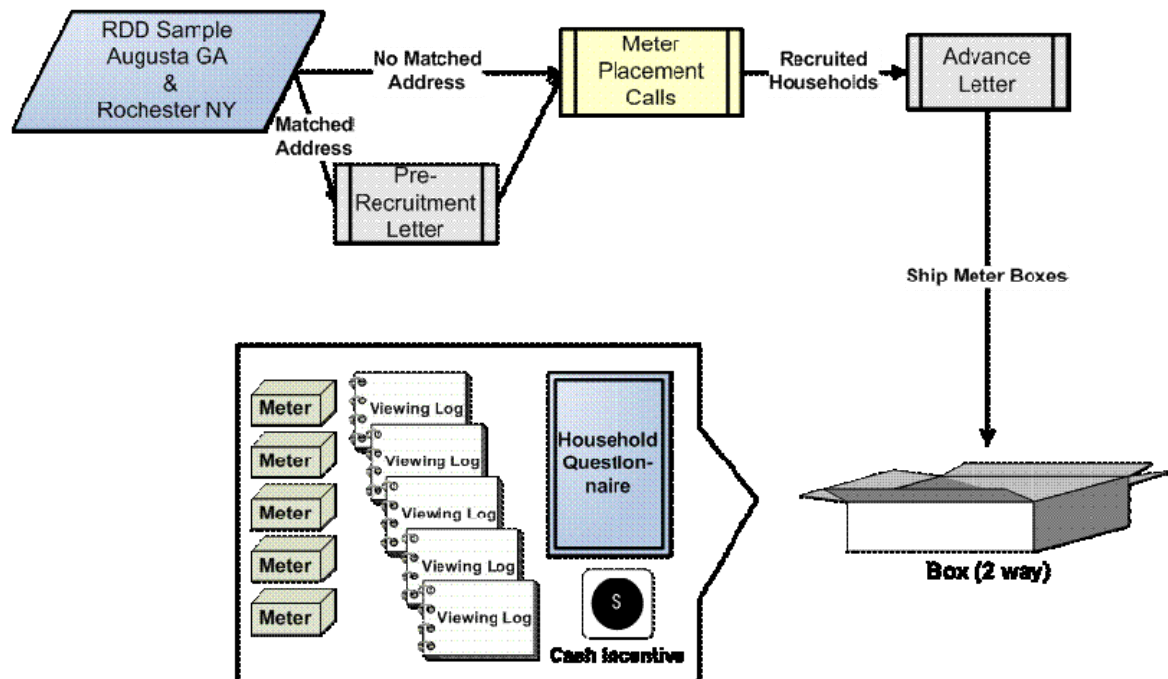


Figure 1. Recruitment and Shipping Process

2.1 Recruitment Techniques

The November study sample was selected with a goal of obtaining 200 “accepting” households (that is, those who agree to participate in the Mailable Meter study) in both the Rochester, New York and Augusta, Georgia designated market areas (DMAs). DMAs are rigidly defined geographical areas used by Nielsen to identify television stations that best reach an area and attract the most viewers. A DMA market consists of all ZIP codes whose largest viewing share is given to stations of that same market area.

Recruitment for the Mailable Meter is conducted via telephone by specially trained Research Interviewers (RIs). For the November 2007 test, a landline random digit dialed (RDD) sampling methodology was used. Given the growing problems associated with landline RDD, future studies involving the Mailable Meter will utilize an address

based sampling (ABS) approach beginning in 2009 (Link, Daily, Shuttles, Yancy, and Bourquin, 2008; Link, Battaglia, Frankel, Osborn, and Mokdad, 2008). Because of the difficulty of reaching certain demographic groups via landline RDD, oversampling was utilized in both markets to gain a proportionate number of accepting households in these hard-to-reach groups. In this instance oversampling included: 1) Age of Householder (AOH) under 35 in both DMAs and 2) Black Householder in Augusta, Georgia. Also, in an attempt to reach more accepting households, refusal conversion was attempted on initial refusals that had not heard the explanation of the survey (that is, “uninformed” refusals).

Although the meters for the Mailable Meter study are less expensive than the traditional Nielsen metering devices, they were still much more costly than the Nielsen TV Ratings Diary, which is mailed to households that refuse or are not contacted in the initial recruitment call if they have a valid address. Therefore, Nielsen only sent Mailable Meters to households that definitely agreed to participate and explicitly understood what they were being asked to do. To accomplish this, the RIs were trained to consider a household to be accepting only if a very clear understanding of the task and commitment to participate was clearly obtained from the respondent. RIs were to “soft close” or thank the respondent and let them know that Nielsen would not be sending them any materials if the respondent did not appear to understand the tasks in full. In other words, all of the households in the study needed to clearly understand and agree to participate.

2.2 Support Structure

Once a household was recruited to participate in the study and sent the materials, Nielsen attempted to keep in regular contact with them to provide a high level of support throughout the study. Up to 8 attempts were made to deliver three reminder calls to participating households – 1) at the beginning of the measurement period, 2) midway through the measurement period, and 3) immediately after the end of the measurement period.

In addition to the reminder phone calls, households were contacted twice more by mail. The first reminder postcard was mailed on November 7th, 2007 to arrive between days nine and twelve of the thirty-five day measurement period. Another reminder mailing (letter) was sent prior to the end of the measurement on November 28th, 2007 to remind the household to return the meter on the appropriate date. The sample management and return process is outlined in the flow chart in Figure 2.

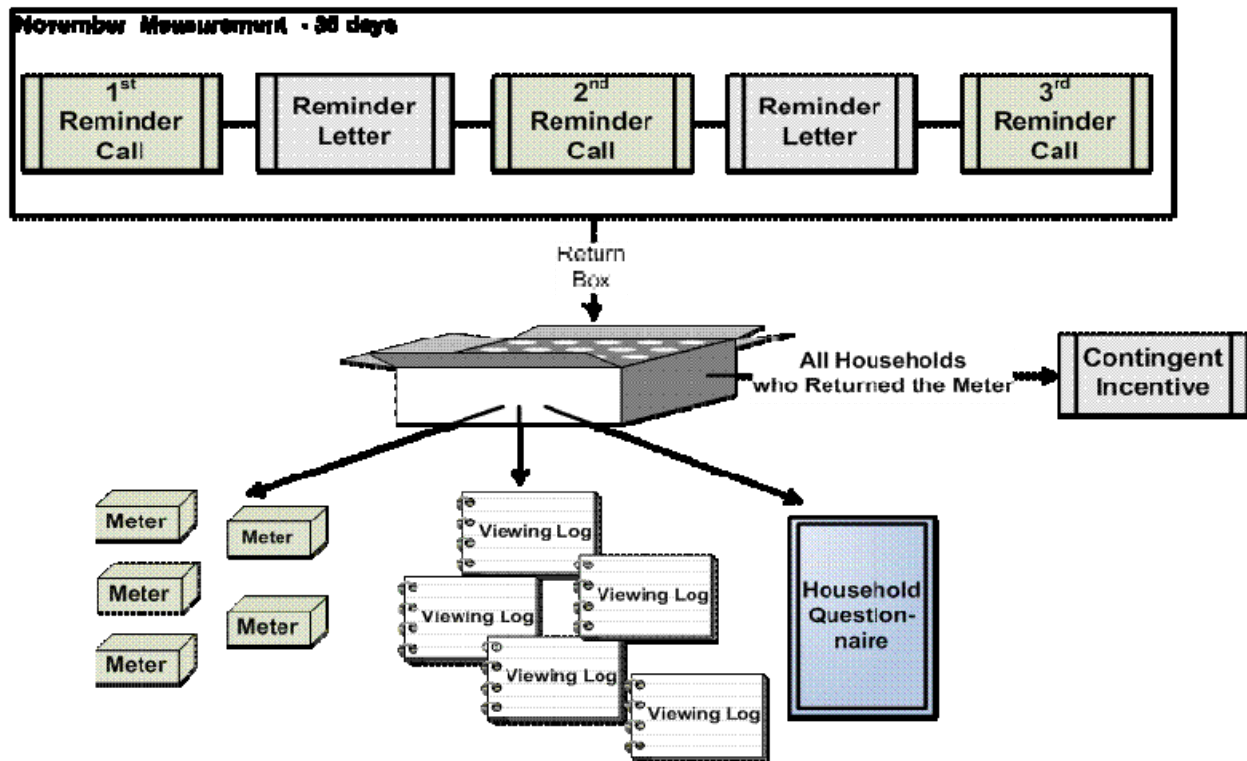


Figure 2. Sample Management and Return Process

2.3 Optimized Incentive Structure

Incentives in amounts ranging from \$50 to \$150 per household were offered for participation (See Table 1). A portion of the incentive (non-contingent) was sent with the materials at the start of the study. Non-contingent monetary incentives are cash money paid to all prospective survey participants before the survey begins. The remainder (contingent) was paid after Nielsen confirmed return receipt of the materials in the form of a check. Contingent incentives are those that are promised to the respondent as a reward for completing the survey.

Previous unpublished Nielsen research had indicated that larger amounts between \$100 and \$200 consistently yielded very high return rates for the Mailable Meter and it was thought that reducing these amounts could help to optimize the use of incentives without significantly hurting return rates. This previous research also indicated that promising an equal or higher contingent incentive amount (when compared to the non-contingent cash incentive) appeared to be more effective than offering a lower contingent amount at the end. Typically Nielsen's Diary service sends a differential incentive to hard-to-reach groups (AOH < 35, Black, and Hispanic) to increase response rates, but that methodology was not included in this first full field test of the Mailable Meter. Nielsen plans to incorporate that into the next round of testing. Households were not paid the contingent incentive unless they returned at least one meter. However, they were paid their contingent incentive if they returned at least one meter even if they did not return the other materials or other meter(s). The checks were mailed out in several waves starting approximately one week after the last day of the measurement period.

Table 1. Incentive Structure

Test Condition	Initial Package Amount	Returned Contingent Checks	Total Incentive
1	\$25	\$25	\$50
2	\$25	\$50	\$75
3	\$25	\$75	\$100
4	\$50	\$50	\$100
5	\$50	\$75	\$125
6	\$50	\$100	\$150

3.0 Results

3.1 Phoning Accepts

Success at the recruitment stage was assessed by calculating a first stage recruitment rate (response rate), based on the number of "accepting" households in the numerator and an estimate of the number of eligible households with one or more televisions in the denominator. This is similar to response rate formula #4 as specified by the American Association for Public Opinion Research (AAPOR 2006). This gives a good indicator of how well the original sample is represented among households included in the study.

As shown on Table 2, there was a nearly 13 percent statistically significant difference between the Diary services recruitment rate and that of the Mailable Meter. The overall Diary recruitment rate for the two DMAs included in this study was 43 percent whereas the overall Mailable Meter recruitment rate was 30 percent.

Table 2. Recruitment Rate - Accepts to Initially Designated Households for Diary versus Mailable Meter

	Accepts				Difference	ID	
	Diary		MM			Diary	MM
Augusta	999	39.8%	209	28.2%	11.6%*	2,509	742
Rochester	1,082	45.6%	192	32.2%	13.5%*	2,372	597
Total	2,081	42.6%	401	29.9%	12.7%*	4,881	1,339

*Significant at the 99% level of confidence

Although the overall recruitment rate was lower than with Dairy, most people were very receptive on the phone if they stayed on the line long enough to hear the explanation of what we were asking them to do. The lower phone recruitment rate was expected as 1) the RI (who was using this script for the first time) needed to obtain clear understanding of the task and informed consent from the respondent 2) the perceived complexity and burden of the task was greater than with Dairy and 3) the recruitment interview did not include differential incentives or assertive persuading.

In examining whether the level of accepts were different for hard to reach demographic groups we examined the percentage of accepts with an under age 35 AOH and with a Black Householder (See Table 3). Obtaining accepts among the under 35 AOH group was more difficult in Rochester than in Augusta and overall more difficult than with Dairy. For the Black group, both DMAs were well represented among the Mailable Meter accepts. In fact, the Universe Estimate was achieved in both DMAs.

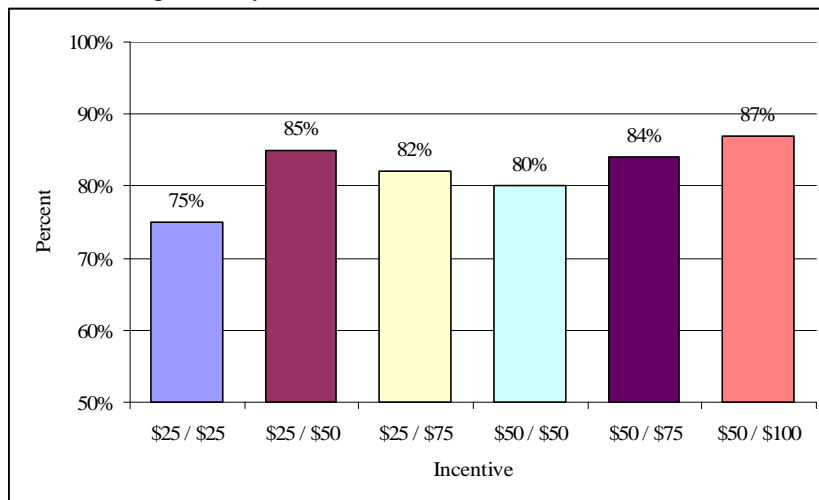
Table 3. Accepts for Hard to Reach Groups

Percent of Accepts with AOH Under 35				
Market	Without Oversample	With Oversample	Diary Without Oversample	Universe Estimate
Augusta, GA	13.6%	18.4%	12.0%	21.0%
Rochester, NY	8.1%	11.9%	12.0%	19.0%
Percent of Accepts with Black Householder				
Market	Without Oversample	With Oversample	Diary Without Oversample	Universe Estimate
Augusta, GA	29.3%	38.2%	30.0%	36.0%
Rochester, NY	12.4%	NA	12.0%	10.0%

3.2 Meter Return Rates

Incentive levels: The overall return rate was 91 percent. In addition, 84 percent of all households returned their materials on-time, where on-time is defined as before 12/19/07 in Augusta and before and 12/21/07 in Rochester (Returns took longer to the Florida production center from New York). Eighty-two percent of participants were compliant in that they returned at least one meter with data and one log with data on-time. Of note, at the \$25 / \$25 incentive level 75 percent of households were compliant. At the \$50 / \$100 incentive level that percentage increased 12 percentage points to 87 percent (See Chart 1). The return rate of meters, logs and questionnaires exceeded expectations. Only 35 households never returned their materials to Nielsen. This high return rate partially mitigated the lower recruitment rate as the Mailable Meter households returned the materials at a much higher rate than did Dairy participants. For these two DMAs, 51 percent of Dairy households who had accepted at the recruitment stage of the survey returned their study materials, which is 40 percentage points lower than with Mailable Meter.

Chart 1: Compliance by Incentive



In response to the follow-up mail survey sent after completion of the meter data collection period, respondents generally reported that they did not see the Mailable Meter task as overly burdensome. Most households reported that: 1) they would like to complete another Nielsen study with meters and logs in the future, 2) thought that the tasks were either as difficult as or easier than expected and 3) the tasks were either as time consuming as or less time consuming than they expected.

3.3 Quality of Return Data

Did returned meter have usable data? In the Nielsen TV Ratings Diary Service “in-tabulation” (or “InTab”) households are those that return a completed diary and pass a stringent editing process to be usable in the Nielsen TV Ratings. Since that editing process is still in development for the Mailable Meter, a preliminary best estimate was used to determine a household level InTab rate by examining the quality of the tuning data that was collected in each meter. The overall InTab rate among households that returned their materials was found to be 85 percent. Some of the meters had no data as the meter may have been unplugged during the test period or not plugged in at all. Also, a computer processing bug caused some meters to be classified as having no data when they actually did have usable data. This caused some households to be classified as not InTab for the whole 35-day period. This data will be reprocessed and the InTab rates will be recalculated. The expectation is that the rate should approach 90 percent. These initial InTab rates indicate that the Mailable Meter, though still in development, currently collects data at an accuracy level approaching that which would be viable in a TV Ratings production scenario.

Did the data in the meters and logs match? Viewing log entries were compared to meter set on / off conditions to produce preliminary meter – log agreement rates. 336 households returned meters and viewing logs that both contained data. Thus the measure of viewing log compliance metric is the rate of agreement between the viewing log and the Mailable Meter at a quarter hour level when the set is on according to the meter. The overall total day meter / log agreement rate for “TV ON” quarter hours for the previously mentioned InTab households was approximately 74 percent. There was some variation by DMA as Rochester had a 77 percent agreement rate whereas Augusta had 72 percent. The meter / log agreement rates by week over the 5 weeks of data collection revealed rates of 71 percent for the first week versus 75 percent for the other four weeks. This indicates that there was no evidence of decline in compliance over the 5 weeks and that there may be a need for a qualification period with Mailable Meter households. Some households may have either plugged their meters in late or did a poor job of complying before we contacted them in the reminder phase of phoning.

3.4 Perceived Task Burden

Another component of household compliance with this combined self-report and electronic measurement methodology is how much participants had difficulty in complying with the tasks they were asked to do. There were two follow-up questionnaire questions that address this concern.

First, respondents were asked if they believed that the task was as easy / difficult as expected. More than nine out of ten households that answered a follow-up questionnaire thought that the task was either easier than or as difficult as they initially expected. Assuming that people would not accept the tasks if they felt that they would be difficult, this reflects well on the ease of the study. However, households with an Age of Householder (AOH) under 35 were more likely to believe that the tasks were harder than expected. Twenty percent of under 35 AOH households responded that the tasks were harder than expected whereas only 7 percent of 35 to 49 AOH households where and 4 percent of 50 or older AOH households was felt this way.

Second, respondents were asked if they believed that the task was as time-consuming as expected. Eighty-five percent of those households that participated and returned the follow-up questionnaire thought that the task was either as time consuming as or less time consuming than they expected based on the initial phone call. Only 13 percent reported that they thought that the task was more time consuming than they expected. Assuming that people would not accept the tasks if they felt that they would be extremely time consuming, this too reflects well on the ease of the study. However, the younger the AOH, the more likely they were to feel that the tasks were more time consuming than expected. Twenty-one percent of households where the AOH was under 35 responded that the tasks were harder than expected whereas 17 percent of households where the AOH was between 35 and 49 and only 9 percent of households where the AOH was 50 or older felt this way.

Of the 298 households that answered the question “What did you think about the frequency of the calls that you received from Nielsen” in the follow-up questionnaire, 75 percent answered “The right amount” and 25 percent answered “Too many”. Thirty-one percent of the respondents in the Rochester DMA answered “too many” whereas 18 percent of the respondents in the Augusta DMA answered “too many”. Nielsen will continue to test various

reminder calling strategies, balancing the need to ensure that participants are given adequate supporting information with the perception that they are being contacted too much or even harassed. Nevertheless, customer service was exceptional in that 99 percent of the 321 households that answered the follow-up questionnaire question, “Were the people you spoke with on the phone during the study courteous / professional?” answered “Yes”.

Results from the follow-up questionnaire also showed that 86 percent of the 316 households that answered the question “Do you remember receiving letters / postcards from Nielsen TV Ratings in regards to this study” answered “Yes”. Thirty-five percent of 281 respondents reported that the letters and postcards they received were not important in helping them complete the study, 45 percent reported that they were somewhat important and only 20 percent reported that they were very important. While these contacts were not generally perceived as important to complete the study, they may have contributed to the household’s perception of the importance of the study.

4.0 Discussion

The information gathered from the follow-up questionnaire indicates that most respondents easily understood the instructions and materials developed for the studies and had little trouble with completing the various steps in the process. Generally speaking, respondents reported that they thought the incentive levels we offered were appropriate. Packaging and receipt of materials and installation of meter was generally well accepted by respondents. The process of receiving and installing the Mailable Meter(s) did not present a problem for most of the participants. Packaging and returning equipment was not a major concern for respondents. Being asked to take a package to a shipping vendor was an effort that most households seemed to anticipate, accept and comply with as part of their participation.

The presence of multiple meters within a household did not appear to cause difficulties for respondents. In fact 86 percent of five meter households had “good” returns.

This study will be followed up with an expanded second full field test of the Mailable Meter which will include one DMA, multiple meters per household (one for each TV up to five), Spanish translations, informed and uninformed refusal conversation, differential incentives to hard-to-reach respondents and additional testing of the overall concept including analysis of log/meter match rates and ratings calculations. In that study the Spanish translations and participation and compliance with the viewing log will be tested. Nielsen is considering the concept of a Mailable Meter panel. Future testing would use an address-based sample (ABS) approach.

Much effort has been made in the recent past to improve and enhance the methodology of the Nielsen TV Ratings Diary service. However, given the explosion in recent years of television viewing options and technologies, a paper diary has become increasingly more challenging for respondents to accurately complete. Hence, there has been a call to replace this methodology with an electronic measurement system. In fact, with an unprecedented initiative coined as Any Time Anywhere Media Measurement (A2M2), Nielsen has made a public commitment to replace the Diary with electronic measurement, most likely a Mailable Meter service, by 2011.

In summary, the idea of combining self reporting with electronic measurement appears to be a concept with the possibility of both easing respondent burden and increasing the accuracy of the data collected in the context of collection of television ratings. Though much research and development is needed to gain confidence in this methodology, other researchers might consider the possibility of such an approach in large ongoing social research studies.

Acknowledgements

We thank the following Nielsen associates, including: Theresa Anderson, Kim Brown, Celia Fernandez, Heather Jordan, Tim Reilly, Jianming Shen, Fred Vereb, Jim Vinyard, Sherri Willms, and in particular the Call Center staff, for their important contributions to project.

References

- American Association for Public Opinion Research (AAPOR). 2006. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*, 4th ed. Ann Arbor, MI: AAPOR.
- Link, M., M. Battaglia, M. Frankel, L. Osborn, and A. Mokdad. 2008. "Comparison of Address-based Sampling versus Random-Digit Dialing for General Population Surveys." *Public Opinion Quarterly* 72, 6-27.
- Link, M., G. Daily, C. Shuttles, C. Bourquin, and T. Yancey. 2008. "Addressing the Cell Phone-Only Problem: Cell Phone Sampling Versus Address Based Sampling." Paper presented at the Annual Conference of the American Association for Public Opinion Research, New Orleans, LA.
- Nielsen Media Research company website:
<http://www.nielsenmedia.com/nc/portal/site/Public/menuitem.55dc65b4a7d5adff3f65936147a062a0/?vgnextoid=cbef0671455bb010VgnVCM100000ac0a260aRCRD> retrieved May 5, 2008
- Schwarz, N. (1999, February). Self-Reports: How the questions shape the answers [Electronic version]. *American Psychologist*. Retrieved April 21, 2008, from <http://sitemaker.umich.edu/norbert.schwarz/self-report>