







COUNCIL FOR RESEARCH EXCELLENCE



Represents advertisers, agencies, networks, cable companies, and station groups





Seeks to advance the knowledge and practice of methodological research

CRE MEMBER COMPANIES



ACCURACY IN RESEARCH

Accurate (accurare - take care)
Research: correct

The Media Rating Council has as its purpose:

To secure ...audience measurement services that are valid, reliable and effective



e: es

VALIDITY AND RELIABILITY

Valid (*validus* - strong)

Research: measures what it claims to measure

Reliable (religare - fasten or bind fast)
Research: produces consistent results



5

AGENDA

Welcome	Richard Zackon Ceril Shagrin
Non-Response Bias Study	Michael Link
Sampling Error Study	Billy McDowell Rick Ducey Steve McGowan
Watcha Watchin' Study	Billy McDowell Michael Link
Audience Participation	Richard Zackon
Break	
Nielsen Response	Paul Donato
Marketplace Panel	Richard Zackon (moderator) Janice Finkel-Greene Billy McDowell Jed Meyer Stacey Schulman
Methodology Panel	George Ivie Matt O'Grady Ceril Shagrin

Wine and Cheese







2012 NON-RESPONSE BIAS STUDY KEY FINDINGS

CRE Sample Quality Committee

Ceril Shagrin

Univision Chair, CRE Chair, CRE Sample Quality Committee Michael Link Nielsen



CRE SAMPLE QUALITY COMMITTEE

Chair

Ceril Shagrin Univision

Members Ann Casey Tribune

Laura Cowan Lin Media

Nancy Gallagher NBCU

George Ivie

Pat Liguori ABC Owned Stations Michael Link Nielsen

Billy McDowell Raycom Media

Beth Rockwood Discovery

Maggie Strickland KVUE

Richard Zackon CRE



8

The sample is the foundation on which research is built. A representative sample is critical to valid, actionable estimates.

The findings from this research are applicable not only to Nielsen local diary measurement but to all research using diary samples.





2009 NON-RESPONSE RESEARCH

- Measured response bias for metered and diary samples >
- Little bias remained after weighting in metered samples >
- Metered samples based on area probability frame with high response rate >
- Bias in diary sample remained even after weighting

Nielsen implemented and addressed based diary frame improving the quality of the frame but dependence on non phone homes to provide contact information reduced overall response rates

No research had been conducted to evaluate the impact of address based samples on response bias.



ISSUES THE RESEARCH WAS INTENDED TO ADDRESS







Improve local diary television measurement



QUESTIONS THE RESEARCH WAS DESIGNED TO ANSWER



What is the difference between responders and non-responders in an address based frame



Can media related equipment ownership be collected from a diary sample



What television programs are being watched in homes without a traditional television set



CONTINUED GROWTH OF CELL PHONE ONLY (CPO) HHS

JULY – DEC 2012

Telephone Status

38.2% HHs did not have landline and have at least 1 wireless telephone

Characteristics of CPO HHs

53.2% adults **age 18-24** 62.1% adults age 25-29 56.7% adults age 30-34

50.5% Hispanic adults **39.0% Non-Hispanic Black** adults

43.9% of Adult living alone 42.2% of Adult(s) + children



Note: Adults are aged 18 and over; children are under age 18. Source: COCNCHS. National Health Interview Survey.



Percentages of adults and children living in households with only wireless telephone service or no telephone service: United States 2003-2012

GROWTH OF CPO HHS FOR KEY DEMOS '09 – '12





CURRENT STATE OF RESPONDER MEASUREMENT

Responder Rates: Nov '08 - Jul '13



Nov 2008 Mar 2009 May 2009 Jul 2009 Nov 2009 Feb 2010 May 2010 Jul 2010 Nov 2010 Feb 2011 May 2011 Jul 2011 Nov 2011 Feb 201 Gross Matched 385,706 399,032 387,758 431,477 421,693 415,386 419,048 416,204 453,897 421,524 424,727 423,102 508,313 481,25 Gross Unmatched 260,367 302,255 322,422 250,162 277,632 298,335 313,553 315,261 321,996 353,336 375,135 376,712 484,485 500,90



121	May 2012	Jul 2012	Nov 2012	Feb 2013	May 2013	Jul 2013
59	500,277	496,199	665,798	603,652	640,736	595,431
03	486,463	474,836	476,222	531,032	487,320	493,109

CURRENT STATE OF RESPONDER MEASUREMENT

Responder Rates: Nov '08 - Jul '13



2012 2009 2010 2011 **42%** to **43% 46%** to **49% 42%** to **51%** 37% to 45%



2012 43% to 47%



→AOH <35 →Black →Hispanic</p>

cre

Jul

2009 NRB STUDY: KEY LEARNING OF RESPONDER HHS

Do Responders differ in their TV viewing from Non-Responders? Yes.

- > **Dayparts**: 11p-2a, 4-8p
- > **Broadcast:** CBS, NBC, Univision
- > Non-broadcast: BET, Cartoon, HBO, MTV



4

Do any differences lead to NRB in TV viewing estimates? Yes.

nonresponse bias

3

Are there differences between **Responder and Non-Responders** on TV viewing correlates?

Yes.

- > **Demo:** HH size, age, presence of children, owner/renter, race
- > Media: Big screen, digital cable, DVR

Yes.

- > Responder status
- adjustments





> All significant differences of viewing estimates in #1 show evidence of

Can NRB be reduced in Nielsen viewing estimates?

> Nielsen poststratification

> Selected TV viewing correlates

NONRESPONSE BIAS STUDY METHODOLOGY PAST TO PRESENT

2009

- **Diary RDD frame** >
- Across Diary DMAs >
- 9K Diary HHs >
- Differential incentive by recruitment > status and responder outcome
- > 4-Mode: Mail, CATI, Web, F2F (full) follow-up

2012

- **Diary ABS frame** >
- 3 local DMAs >
- > 20K+ Responder HHs
- >
- > follow-up



Standard incentive across sample Dual-Mode: Mail, F2F (partial)

2012 NRB STUDY METHODOLOGY AND RESPONSE RATE

JULY - DEC 2012

 Mail Qn Return Rate

 by Responder Outcome across 3 DMAs

 76% to 78% Responder HHs

 (n=410-846)

 60% to 67% No Good HHs

 (n=61-152)

 23% to 28% No Return HHs

 (n=2,430-2,987)

F2F Interview Completion Rate

by Responder Outcome across 3 DMAs

45% to **59%** Responder HHs (n=14-22)

0% to 33% No Good HHs (n=3-6)

45% to **59%** No Return HHs (n=191-196)





2012 NRB STUDY: ANALYTICAL OBJECTIVES





Responder vs. Non-Responder of TV HHs

- > Demo and geography
- > Media access and usage
- > Viewing measures



- Standard responder weights
- > CRE responder weights
- > Auxiliary data







Zero-TV HHs

- > Demo and geography
- > Media access and usage
- > Viewing measures

KEY LEARNING OR RESPONDER AND NON-RESPONDER HHS

Responder vs. Non-Responder of TV HHs Differences

HOH Characteristics

> Age, race, Hispanic ethnicity, work status, education

Media Usage

> Game system, MP3 player, cell phone activities

Viewing Behavior

- > Early morning and prime time daypart
- > Broadcast, nonbroadcast stations

NRB and **Weighting Method**

Non-Response Bias

> NRB detected for all differences by dayparts and stations

Weighting Method

- > Nielsen PS weights reduced some bias
- > CRE responder weights closer to Nielsen mediarelated UEs (than Nielsen PS weights)

Auxiliary data yielded little insights

> Game system, Internet access

> Non-TV viewing on desktop/online



Zero-TV HHs Differences

HH Characteristics > Age, language, income

Media usage

0-TV HH Viewing Behavior

> OOH viewing

RESPONDER VS. NON-RESPONDER HOH/HH PROFILE

Responder HOH Characteristics

- > More likely over age 50
- > More likely of White race
- > Higher educated (bachelor's or higher)
- > Less likely in work force

Responder HH Characteristics

- > Less likely to have children in the HH
- More likely to own their home and live there for 10+ years





HOH AGE: HIGHER RESPONDER AGE 50+ ACROSS ALL 3 DMAS





HOH RACE/HISPANIC ETHNICITY: HIGHER NON-RESPONDER NON-WHITE & HISPANIC ACROSS ALL 3 DMAS





RESPONDER VS. NON-RESPONDER PHONE/MEDIA PROFILE

Landline/Cell Phone of Responder

- > HHs more likely to have landline
- > HHs more likely to own only one cell phone

Media-Related for Responder

- > Less likely to own any gaming system
- Less likely to own MP3 player (iPod, Zune)
- Less likely to access Internet on cell phone
- > Less likely subscribe to video svc on cell phone





Different

INTERNET AND VIDEO VIA CELL

MP3 PLAYER

GAME SYSTEM

LANDLINE/CELL

LANDLINE/CELL PHONE: HIGHER NON-RESPONDER HHS WITH **NO LANDLINE AND RESPONDER HHS WITH ONLY 1 CELL PHONE**



Responder

Non-Responder



MEDIA OWNERSHIP & USAGE: HIGHER NON-RESPONDER **HHS WITH MEDIA DEVICES & USAGE**



Responder

Non-Responder

cre

RESPONDER VS. NON-RESPONDER VIEWING PROFILE

Responder Viewing by Dayparts

- > Higher person viewing early morning and prime time across all 3 DMAs
- > Higher HH viewing differ by dayparts and by DMA

Responder Viewing by Stations

- > Lower person viewing across non-broadcast stations
- > Higher person viewing across broadcast stations except Univision

Similar
FOX, A&E, CNN, TNT, USA
PERSON 11P-2A





PERSON VIEWING BY DAYPART: HIGHER DIARIES EARLY **MORNING, PRIMETIME & OVERALL ACROSS ALL 3 DMAS**

NRB DETECTED FOR THESE DAYPARTS WITH PERSON VIEWING OVER-ESTIMATED FOR DIARIES (P<0.01)





HH VIEWING BY DAYPART: HIGHER RESPONDER **EARLY MORNING, PRIMETIME & OVERALL**

NRB DETECTED FOR THESE DAYPARTS WITH HH VIEWING OVER-ESTIMATED FOR RESPONDER (P<0.01)





PERSON VIEWING BY NON-BROADCAST STATIONS: HIGHER NON-RESPONDER ACROSS ALL STATIONS



- 7 out of 11 non-broadcast stations consistently have higher percentage of persons viewing from non-Responder compared to Responder (p<0.05)
- Cartoon Network has higher percentage of persons viewing from non-Responder across all 3 DMAs (p<0.05)
- Nonresponse bias detected for these stations with viewing under-estimated for Responder



Dallas





PERSON VIEWING BY BROADCAST STATIONS: HIGHER RESPONDER FOR ENGLISH-LANGUAGE STATIONS



- 3 out of 4 English-language broadcast stations have higher percentage of persons viewing from Diary compared to non-Responder (p<0.05)
- NBC has higher Responder/persons viewing across all 3 DMAs (p=0.00)
- Univision, on the contrary, has higher non-Responder/persons viewing in Dallas (p=0.00) and Albuquerque (p=0.07)
- NRB detected for these stations with viewing over-estimated for Diaries except for Univision

Red stations indicate the same stations with viewing differences from 2009 NRB study

cre

Dallas



UNIVISION *

DALLAS LPM VS METER/ADJUSTED DIARY RATINGS **BROADCAST COMPARISON BY DAYPARTS AND DEMO**



Univision: M-F 12A-12A





NBC: M-F 8P-11P





Univision: M-F 8P-11P

DALLAS METER/ADJUSTED DIARY VS DIARY ONLY RATINGS **NON-BROADCAST COMPARISON BY DAYPARTS AND DEMO**









NBC: M-F 8P-11P





Univision: M-F 12A-12A

Univision: M-F 8P-11P

NIELSEN'S POST-STRATIFICATION WEIGHTING REDUCED BIAS BUT SIG DIFF REMAIN



- Significant differences of HH viewing (12a-12a) of Diaries and total sample remain for all 3 DMAs (after PS weighting) >
- Significant differences of person viewing (12a-12a) of Diaries and total sample remain for all DMAs but Dallas (after PS weighting)


NIELSEN'S POST-STRATIFICATION WEIGHTING REDUCED BIAS BUT SIG DIFF REMAIN





- Significant differences of broadcast viewing (NBC) of Diaries and total sample remain for all DMAs except Dallas > (after PS weighting)
- Significant differences of non-broadcast viewing (Cartoon) of Diaries and total sample remain for Albuquerque and Paducah (after PS weighting)



Cartoon

CRE ESTIMATE WITH RESPONDER WEIGHT CLOSER TO NIELSEN ESTIMATE FOR "CABLE AND/OR ADS"



CRE1 weighted by responder weight and CRE2 weighted by Nielsen post-stratification weight >



CRE ESTIMATE WITH RESPONDER WEIGHT CLOSER TO **NIELSEN ESTIMATE FOR "CABLE AND/OR ADS"**



CRE1 weighted by responder weight and CRE2 weighted by Nielsen post-stratification weight >





Andy Peytchev, PhD Senior Survey Methodologist

RTI WEIGHTING ADJUSTMENT RESULTS

Objectives and Approaches

- Replicate weighting results by Nielsen and suggest possible other weighting approaches
- Components of RTI's weights
 - Selection probability
 - Nonresponse within selected sample
 - Poststratification for nonresponse and coverage

Results and Next Steps

- bias remains
- > Evaluating other potential approaches for weighting adjustment
 - Person/HH level data from commercial sources



> Similar to Nielsen's results, RTI weights reduced nonresponse bias but most of the

Cross-classification of characteristics

RTI SUMMARY



This nonresponse bias study, repeated over time, is a commendable effort from CRE and Nielsen



The analyses of the study are sound



There may be little benefit to be gained from improving poststratification adjustments, but analytic domains may benefit from more detailed breakdown of population totals



We are currently evaluating different changes to the nonresponse weight component, which will be included in our final report



ZERO-TV HOUSEHOLDS





ZERO-TV HOUSEHOLDS: HOH CHARACTERISTICS





ZERO-TV HOUSEHOLDS: MEDIA DEVICE OWNERSHIP/USAGE



63% Computers with hi-speed internet



4% Game system with hi-speed internet



26% Computers without hi-speed internet



0% Game system without hi-speed internet





96% No game system





41% Use internet on cell phone



41% Do not use internet on cell phone





ZERO-TV HOUSEHOLDS: VIEWING BEHAVIOR



cre

*Viewing by location reported if HOH has OOH viewing (no location capture at viewing level)

Viewing by Location* n=15 HOHs 12% se's home Viewing by Location*



DIFFERENCES OF TV VS. ZERO-TV HHS (GENERAL OBSERVATIONS FROM 27 HHS...)

0-TV HOH Demo

- Younger (higher proportion of AOH<35)</p>
- Speaks other than English at home
- Lower HH income (higher proportion of <\$40K)</p>

0-TV HH Media Usage

- > Less likely to have game system
- Less likely to have computer with high speed Internet





0-TV Person Viewing

Most viewing reported on TV

Most non-TV viewing on desktop (device) and from Website (source)

 A third is OOH viewing (most at someone else's home)

IMPLICATIONS FOR FUTURE STATE: EFFECT OF CPO HH CHARACTERISTICS

CPO HHs with Children

(from 21% in '09 to 45% in '12)

- > Non-Diaries has 10-14% higher proportion of households with children than Diaries for both **NRB** studies
- > Non-Diaries have higher viewing estimates for Cartoon Network than **Diaries for both NRB** studies

CPO HHs with **Hispanics**

(from 28% in '09 to 51% in '12)

- > Non-Diaries has 6% higher proportion of Hispanic HHs in Albuquerque & Dallas for '12 NRB study
- > Non-Diaries has higher viewing estimates for Univision than Diaries consistently for both **NRB** studies





CPO HHs w/ Young Adults under 35*

(from 34% in '09 to 57% in '12)

> Non-Diaries has 3-10% higher proportion of AOH<35 for both NRB studies

> Non-Diaries has higher viewing for BET, Cartoon, MTV cited for both NRB studies

IMPLICATIONS OF FUTURE STATE: EFFECT OF UNMEASURED HH CHARACTERISTICS

Zero-TV HHs

(2012 NRB Study)

- Demos: Young adults (under 35), speak other than English at home, lower HH income
- Viewing Behavior:
 Single-person viewing,
 TV viewing at someone
 else's home, non-TV
 viewing on online

Non-Responder HHs

(2009 & 2012 NRB Study)

- Demos: Younger HOH (under 35 & 35-49), non-White, Hispanic, larger HH size (with children)
- Viewing Behavior:
 Non-broadcast viewing,
 Spanish-speaking
 broadcast viewing
 (Univision)



Mobile TV Viewers

("TV Untethered: Measuring the Shifting Screen")

 Demos: Young adults (mean age 35), African-American, English-dominant Hispanic, larger HH size

> Viewing Behavior:

Mobile viewing in the home, online, during daytime, primetime & late fringe

IMPLICATIONS FOR FUTURE STATE: FURTHER WEIGHTING CONSIDERATIONS

Nielsen/ RTI Weighting Method

- Alternate production Responder weighting process was used by Nielsen and potentially impacted the effectiveness for NRB adjustment
- Nielsen/RTI weights reduced bias but not significantly

CRE Responder Weights

- Special weighting variables were created for the NRB study using homes completing the NRB questionnaire
- CRE responder weights are closer to Nielsen media-related UE.
 Nielsen PS weights more effective for reducing NRB.



Auxiliary Data Sources

 Multivariate analysis was used to detect any covariates from 2K+ auxiliary vars. to adjust for NRB

 Significant variables selected as covariates from PRIZM and POI sources showed mixed pattern for segmentation and/or geographic landmarks













TREND ANALYSIS OF RELATIVE ERROR OF RATINGS IN NIELSEN LOCAL TV DIARY MARKETS

CRE Local Measurement Committee

Billy McDowell

Raycom Media Chair, CRE Local Measurement Committee **Rick Ducey Steve McGowan** BIA/Kelsey

CRE LOCAL MEASUREMENT COMMITTEE

Chair **Billy McDowell Raycom Media**

Members Ann Casey Tribune

Joanne Burns 20th Television

Laura Cowan Lin Media

Janice Finkel-Greene Magna Global

Bruce Hoynoski Nielsen

Mark Kaline Kimberly Clark

Pat Liguori ABC Stations

John McMorrow CoxReps

Jed Meyer Annalect

Keenan Pendergrass Cox Media

Rick Pike Intermedia Adv

Matt Ross NBCU

Univision

KVUE

NAB

CRE



- **Julie Russell** AdCo Advertising
- **Ceril Shagrin**
- **Maggie Strickland**
- **Sharon Warden**
- **Richard Zackon**

EXECUTIVE SUMMARY

- The Local TV Diary is the currency for 30% of all television viewing.
- Viewing data continue to play critical role in local sales efforts but the diary-based ratings show the effects of audience fragmentation and declining response rates.
- The number of in-tab diaries declined by nearly 16% in the 31 markets analyzed from 2001 to 2012.
- When smaller sample sizes are combined with lower ratings caused by audience fragmentation since 2001, Relative Error rates rose at a higher rate, usually exceeding the +10% industry business rule.
- Relative Error was somewhat smaller for affiliates of the four largest broadcast networks compared to the strongest Independent stations in the 31 markets.



Because of generally larger ratings in prime time, they are more stable and more frequently within the sales benchmark, but always less than one-half of the time. The acceptable RE scores are smaller for all the Persons demo breaks, especially for P18-34s.



OBJECTIVES

- > Purpose and Background
- > Methods
- > Issues with Local TV Diary
- Industry Expert Interviews: Buy and Sell Sides
- > Trend Analyses of Relative Standard Error
 - May Sweeps 2001, 2003, 2005, 2007, 2009, 2011, 2012
- Insights from the Interviews and Analyses
- > Looking to the Future





PURPOSE AND BACKGROUND



Council for Research Excellence's Local Measurement Committee issued RFP to obtain a, "thorough analysis of two sets of data Nielsen has provided on eleven years' trend of Relative Errors in 31 diary markets and posted advertising schedules."

> Data Set #1

Seven May sweep periods (2001, 2003, 2005, 2007, 2009, 2011 and 2012)

> Data Set #2

Ratings, Standard Error and Relative Error Estimates



CRE retained BIA/Kelsey for this project.

- > BIA/Kelsey is a research, advisory and consulting firm that has served the broadcast industry since 1983.
- > Principal consultants have deep experience in various aspects of television audience measurement.



The overall goal of the project is to examine potential instability in ratings as evidence by relative error.



METHODS

To establish context and mine "institutional knowledge" around the local television diary in diary-only markets we conducted a series of expert interviews.

> Buyers, Seller, Nielsen interviews

Nielsen provided six May sweep data from 2001-2012.

Analyses included:

- > Relative Error (RE) range for All Stations, Affiliates, Households, Demos
- > HUTs and RE by Daypart and Demos
- > RE by Effective Sample Size (ESS) by Daypart and Demos

In-Tab and Effective Sample Size

Included breakouts of RE error in the 0-10% range since this is a key factor impacting the ratings "currency" value in the marketplace.





QUICK PRIMER ON RELEVANT TERMS

EFFECTIVE SAMPLE SIZE

The size of random sample that would provide the same standard error as the actual sampling plan on which a survey result is based.

SAMPLE ERROR

Statistical measure of the possible deviation of a sample estimate from the true population value, assuming the sample to be representative of the population from which it has been drawn. The sample error is normally expressed as a margin of difference either side of the reported value within specified confidence limits (i.e. "there is an x% probability that the true population value lies within y units either side of the sample estimate"). Sample error is wholly distinct and not to be confused with sample bias, for which no parametric statistical assumptions can be made.

STANDARD ERROR

Standard deviation of the sample error distribution of sample estimates. 1.96 standard errors denotes the upper and lower bound margins of sample error that correspond with 95% confidence limits.



ISSUES WITH THE LOCAL TV DIARY

Legacy Methodology

Measurement And Estimation Issues

- Lower average ratings; smaller effective sample sizes. >
- Lower response rate; reliability and validity of respondent diary entries. >
- Measurement limited to sweeps weeks. >
- Forecasting models use diary data to predict ratings up to 12 months out. >

Consumer "Viewing" Has Multiple Meanings

- Linear TV channels; DVR; streaming; VOD; Mobile; Connected TVs; > Tablets; Game consoles; Live versus time-shifted ("C7"); more choices, etc.
- Too many choices for paper diary. >

EMRC Accreditation Withdrawn In 2010

Diaries unreliable as form of measurement.

Industry

cre

- Ill-informed about local ratings stability issues. >
- Local held to higher standard than national. >
- Focus on small cells versus broader demo/daypart cells. >



NIELSEN RATINGS / NORFOLK, VA.

Network	Station	Rating
ABC	WVEC	8.3
NBC	WAVY	7.0
UPN	WGNT*	4.2
CBS	WTKR	4,1
WB	WTVZ*	2.2
Fox	WVBT*	1,5
Pax	WPXV*	0.8
ABC	WVEC	8.3
NBC	WAVY	7.7
UPN	WGNT*	4.2
CBS	WTKR	4.1
WEI	WTVZ'	3.2
Fox	WVBT*	1.0
Pax	WPXV*	0.8
NBC	WAVY	8.9
ABC	WVEC	8.4
UPN	WGNT*	5.3
CBS	WTKR	5,0
WB	WTVZ*	3.2
Fax	WVBT*	1.7
Pax	WPXV*	0.9
Fax	WVBT	3.8
NBC	WAVY	11.0
ABC	WVEC	7.8
CBS	WTKR	6.5
UPN	WGNT*	4.0
WB	WTVZ*	2.1
Fox	WVBT*	1.3
Pax	WPXV*	0.3

"Non-news programming: Source: Nielaen Media Rassarch, May 2003

ECONOMICS OF LOCAL TV DIARY

~DMA 57+ use only diary data

> 30% of the U.S. television viewing households are measured by local television diary survey research methods.

12% of TV ad spend based on diaries.

Biggest issue is cost

- > Makes local television seem less profitable.
- > Stations moving to STB-based measurement as alternative.

Economics make it hard to justify remedies such as increased effective sample size, electronic measurement.

Local stations argue they are held to unreasonable level of accountability.

> They get penalized for under delivery but not credit for over delivery.

One buyer gave an example

- > 10% of a \$20 million budget spent in local TV is \$2 million.
- > With the 10% rule in play, that means up to \$2 million at risk in this example.



I understand that anything they [Nielsen] do to improve diaries is going to increase the cost associated with it... They might see an exodus of clients."

Justin Lewis Research Director Fisher Communications

ALTERNATIVES TO THE LOCAL TELEVISION DIARY

Set Top Box (STB) Measurement

- > Rentrak, Nielsen
- > Large "samples" but not statistical samples
- > Household data only

Nielsen's acquisition of Arbitron and Personal Portable Meter (PPM)

- > Might have some applicability
- > Could retire diary altogether

Nielsen Code Readers

- > Local People Meters + STB + Code Readers
- > Goal is to increase sample size; ratings stability

Coalition for Innovative Media Measurement

> Working with Nielsen on "electronic diary" using PC, smartphone, tablet

Social Media

Increasingly used as indicative of viewing but cannot substitute for measured viewing.

Adjacent Meter Market

> Used as imperfect proxy for guidance.



Despite the desire for a perfect data source for local TV ratings, the industry must work with the limited options available. Marketers should understand the limitations of the data and look forward to improved measurement

Julie Pahutski Empower Decision Sciences

WHAT ARE THE BASIC ISSUES WITH LOCAL TV DIARIES?

The expectation is that dairy can be used to forecast up to 12 months out.

> Using May or July books to forecast period until November data are available.

There is no ability to self-correct these forecasts.

> No independent source data.

Stations are faced with the need to offer ratings guarantees.

> Even for programs that may or may not run; or not even being broadcast during sweeps period used for forecast.

Sample sizes are too small for number of viewing sources and choices.

- > Audience fragmentation
- > Decreased average ratings

The economics don't justify increased sample or electronic measurement.

Hybrid models, such as using STB data, introduce problems.





These issues could be resolved by explaining the obstacles."

> Janice Finkel-Greene Magna Global

WHO'S IMPACTED?



10



Hoak Media LLC

507 TV STATIONS

LOCAL CABLE IN A DAIRY-ONLY MARKET





COMMENTS ON THE LOCAL TV DIARY SYSTEM

"You know what, these days we buy video impressions. Our clients want local video impressions and those can come from local TV but also from streaming, Digital Out Of Home video screens, etc. We need to feel as confident in the TV data as we do with the digital video data."

"The diary measurement system may not be the best, but it's the only hard data we have."

"If the Code Reader doesn't advance and we're left only with diaries that have limited accuracy and stations continue to drop [the Nielsen] service; that would cause us to reconsider."

"10% [accepted relative error] is not good enough, but we have to use something." "We won't abandon buying local television because of diary issues, there will always be a need for local television."

"We look at STB data from Rentrak and stations' own studies. But we cannot buy off these data. They must provide additional inputs."

THE 10% BUSINESS RULE FOR RELATIVE ERROR

- > The industry, both buy and sell sides, accepts that Ratings should be viewed as point estimates with a margin of error to be associated with these estimates.
- > By convention, the buying and selling on local television diary data accepts forecast Ratings bought and sold as equal to Ratings in the posting process if they are within $\frac{1}{-10\%}$ of the Relative Standard Error, or the Standard Error as a Percentage of the Estimate Rating.
- > Ratings delivered less than 10% of forecast trigger additional negotiation between buyers and sellers in successive buying rounds.
- > Remedies may include pricing discounts or "make-goods" (i.e., Alternative Delivery Units (ADUs) or spots provided to compensate for under-delivery of promised audiences.

Accepted 10% 5.4% Confidence Interval



EXAMPLE

Ratings Estimate Time/Demo Rating estimated to be



+/- relative standard error of 10% or 0.54%.

5.4% +/- 0.54% = 4.86% to 5.94%

as "accepted" Rating

Ratings less than

are judged to be less than promised

HOW CAN THE SITUATION BE IMPROVED?

"Look only at aggregated ratings, not by demo."

"Code reader."

"Well, a number of stations will use data from a nearby, comparable metered market and use daypart and demo data to adjust what we're seeing in the data from the diary-only markets. This practice has some acceptance." "What the industry needs is some sort of statement regarding accepting the [local television diary] data for what they are."

"Use Set Top Box data and complement these data with demo data." "We have business rules that do not imply an understanding of statistics."

> "Using mobile devices like smart phones to help measure viewing. Everyone has them, it can't be that expensive."

USING THE LOCAL TV DIARY TO BUY AND SELL

Ratings and Confidence Intervals

- > Nielsen does diary measurement in four sweeps periods.
- > The data in study examine sweeps data for diary-only markets in May 2001, 2003, 2005, 2007, 2009, 2011 and 2012.
- > Ratings are point estimates that have associated error due to the sampling process, i.e., standard error.
- > Properly understood, a ratings point is an estimate that has an associated confidence interval that is estimated by the standard error.



RELATIVE ERROR FINDINGS: ALL STATIONS

- > In Total Day, the relative error of all diary-only household ratings falls within the acceptable range just 11.3% of the time, meaning nearly all fluctuation can be attributed to statistical error.
- > Because of generally larger ratings in prime time, they are more often within the sales benchmark, but still only 26% of the time.
- > During the traditional local news windows (weekdays 6-7pm and 11-11:30pm) ratings for all stations are within the sales range 18.1% and 20.7% of the time, respectively.



RELATIVE ERROR RANGE ANALYSES





		Mon	Tue	Wed	Thu	Fri	Sat	Sun
DAYPARTS	6:00A-6:00A							
	6:00A-7:00A							
	7:00A-4:00P							
	8:00P-11:00P							
	11:00-11:30P							
	11:30P-1:00A							

DEMOS 18-34 18-49

TREND DATA INCLUDED FROM MAY OF EACH YEAR

2001	2003	2005	2007	2009
MAY	MAY	MAY	MAY	MAY







RELATIVE ERROR TREND ANALYSIS FINDINGS: NETWORK AFFILIATES

RELATIVE ERROR RANGE 93.0% high 57.1% M-F primetime day part (8P-11P) **Total Week Relative Error** falls inside the 0-10% "accepted 45.3% low M-F-4P daypart range" an average of 57.1% of the time

- > As the large group of stations is subdivided into smaller components the data become less stable.
- In Total Day, the relative error of all diary-only household ratings of affiliates of the four largest broadcast > networks were nearly all beyond the acceptable relative error range. The largest incidence of acceptance was 3.9% for CBS affiliates. None of the ratings for Fox affiliates were within the +10% range, meaning all fluctuation can be attributed to statistical error alone.
- Prime time ratings are more often within the sales benchmark, ranging from 26% (CBS) to 10.1% (Fox). >
- During the traditional local news windows (weekdays 6-7pm and 11-11:30pm) acceptable relative error for ABC, NBC and CBS affiliates were somewhere between the numbers from prime and Total Day, ranging from 10-18%.



ZERO CELLS

No zero cells reported
RELATIVE ERROR TREND ANALYSIS FINDINGS: "STRONGEST INDEPENDENT STATIONS"



- > As the large group of stations is subdivided into smaller components the data become less stable, meaning Relative Error increases.
- > Nielsen identified the single largest independent station in each of the 31 markets and isolated their Relative Error ranges.
- > None of the ratings for these independent stations were within the +10% range, meaning all fluctuation can be attributed to statistical error alone.



ZERO CELLS Range from 21.1% to 79.5%

MAY 2012 RELATIVE ERROR BY DEMOS: HOUSEHOLDS

RELATIVE ERROR

23.9% **Total Week Relative Error fell** within the 0-10% "accepted range" 23.9% of the time



- > Across all 31 markets Nielsen produced Household Relative Error ranges.
- In Total Day, the Relative Error of all diary-only household ratings > falls within the acceptable range about 24% of the time, meaning three-fourths of all fluctuation can be attributed to statistical error.
- Because of generally larger ratings in prime time, they are more > often within the sales benchmark, but still less than one-half of the time (46%).
- During the traditional local news windows (weekdays 6-7pm and 11-11:30pm) household ratings are virtually the same: within the sales range 39% of the time.







MAY 2012 RELATIVE ERROR BY DEMOS: HOUSEHOLDS

RELATIVE ERROR

1.4%

Total Week Relative Error fell within the 0-10% "accepted range" 1.4% of the time



RELATIVE ERROR

6.7%

Total Week Relative Error fell within the 0-10% "accepted range" 6.7% of the time



RELATIVE ERROR

13.4%

Total Week Relative Error fell within the 0-10% "accepted range" 13.4% of the time



25-54

18-3

to

28.4%

RELATIVE ERROR FINDINGS: PERSONS 18-54

- > As the diary population is subdivided into smaller age components (18-34, 18-49 and 25-54) the data become less stable than for households.
- > The 25-54 year old demographic comprises the largest number of respondents in the diary sample and, according, they produced the largest incidence "acceptable" ratings based on the sales threshold of +10%. Even so, in prime time when ratings are highest and relative error the lowest, the RE is within the 10% rating about one-fourth of the time (28%).
- > Relative error for 18-34 year olds yield the smallest acceptable RE numbers, exceeding the 10% mark virtually all the time across. They only fell below the threshold 1.4% for Total Day and 13% in prime time.

> 18-49 year olds fall between the two: 7% in Total Day and 18% in prime.



RE VERSUS EFFECTIVE SAMPLE SIZE MON-SUN 24 HOURS, 18-49

- The number of 18-49 diary respondents has been steady the past three surveys, around 250 > but that is more than a 10% decline since 2007 alone and nearly 30% since 2001.
- Total Day Relative Error reached its high mark in 2012, up by 25% since 2001 >





LOOKING TO THE FUTURE

- The long-established +10% threshold for audience delivery guarantees are outdated given the diary's increasing relative error rates
- The diary's sweeps measurement frequency (4x/year) no longer reflects today's year-round programming environment and makes it difficult to predict and report on programs that fall between the February, May, July and November measurement periods
- Solutions for reducing relative error rates possess financial barriers, lack the diary's granularity and will be challenged by technology trends as consumers find more ways to watch more content
- The buying and selling communities are still evaluating the viability of using Set Top Box data sources for transaction purposes



Some media buyers could consider replacing local TV with other media as TV diary accuracy erodes







The Watcha Watchin'? Experience

Billy McDowell, Raycom Media | Michael Link, Nielsen

Whatcha Watchin'? Overview



Capture viewing on any viewing platform

User Registration

Do you own any of the following devices? Check all that apply.



Reporting platform

Device choice



Raycom Media - University of Alabama - Nielsen Company

Key Learning

- Majority respondents used the website version and stayed with a single mode/device
- Note: For this study the mobile app was only available for Android phones (no iOS version was available)

Reporting Behavior by platform

Number of viewing entries



Average Number of Viewing Entries Reported per Day

Raycom Media · University of Alabama · Nielsen Company

Key Learning

- Those who used a
- smartphone made, on
- average, more program
- entries than did those
- who used online/web
- (2.1 vs 1.7)
- This finding holds across
- each day of the week

Reporting behavior by platform

Live reporting vs. Retrospective reporting



*Live report: reported the viewing the same day as watched. Retrospective report: reported the viewing on a different day. Key Learning Those who used a smartphone made a higher percentage of entries on the same day they viewed the program (73% vs 55%) Significant implications for data quality as retrospective reports have a higher potential for recall bias

Viewing behavior by platform

84%

Web

Online viewing & Out of home viewing

50%

40%

30% 20%

10% 0% 80%

Smartphone



Key Learning

Out of Home

In Home

Raycom Media · University of Alabama · Nielsen Company

programming content online as compared to those using smartphones Those using smartphones were somewhat more likely to report out-ofhome viewing than were those making Web entries

A slightly higher percentage of those making Web-based entries reported viewing

Experience Survey

Whatcha Watchin?



Raycom Media - University of Alabama - Nielsen Company

Experience Survey

How was your overall experience as a participant?



Media Usage Survey



Viewing behavior

Type of Program Format Reported



Raycom Media · University of Alabama · Nielsen Company

Key Learning Majority (65%) of viewing was reported as traditional TV

Nearly one-quarter (24%) of viewing was reported as program content accessed online

Confidential

Media Usage

Do you follow news, sports or weather on the web sites below?



Raycom Media - University of Alabama - Nielsen Company





Media Usage

63% of respondents who reported turning to Twitter for breaking news also follow a local TV station or personality on Twitter.

27% of respondents who reported station.





Raycom Media - University of Alabama - Nielsen Company

turning to Facebook for breaking news are also Facebook friends with a local TV

Confidential 93

Questions





MARSING BAR

RULES

In groups of 3-5, draft a brief research question (25 words or fewer) the answer to which would likely inform measurement methodology



Nielsen Response

Paul Donato EVP Chief Research Officer nielsen



97

MARKETPLACE PANEL







METHODOLOGY PANEL





THANKS FOR YOUR ATTENTION!

