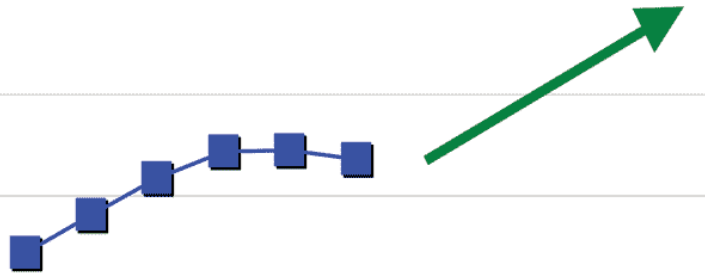


# AUDIENCE 2010



## Reinvigorating Public Radio's Public Service & Public Support

### INTERIM REPORT 3

## RELIABILITY & INTEGRITY

### AUDIENCE 2010

is a project of

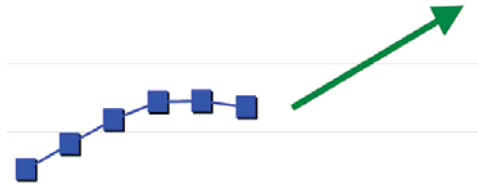
Walrus Research  
&  
AudiGraphics, Inc.

for the

Radio Research Consortium

February 28, 2006

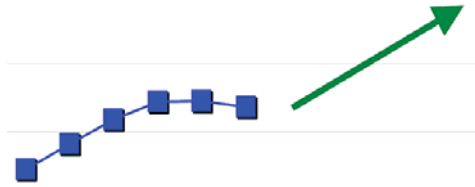
Copyright © 2006 All Rights Reserved



All audience estimates in this analysis are based on radio listening as recorded and copyrighted by Arbitron.

Thanks to Dr. Ed Cohen of Arbitron for providing Arbitron's response, return, and consent rate data shown on pages 6 and 8.

This report is copyright © 2006 by the Radio Research Consortium, Inc., Walrus Research, Inc., and AudiGraphics, Inc. All rights are reserved.



## Findings

Arbitron estimates indicate that public radio lost audience between 2003 and 2005, and that the momentum which powered its long run of national audience growth is gone.

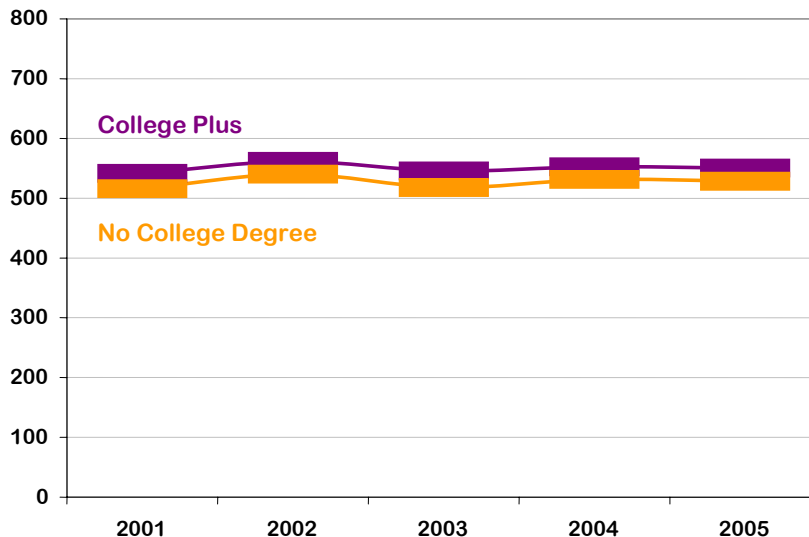
Might deficiencies in Arbitron's sampling, response rates, or method of measurement be causing *apparent* rather than real declines?

We dug into Arbitron's methodology, with particular reference to public radio listeners.

**We find Arbitron's national estimates reliable. Public radio's national loss of audience momentum is real.**

**The same objective criteria call into question the reliability and integrity of a significant body of research on which public broadcasters now base decisions.**

**Average PPDV**  
Persons 18+  
Arbitron Spring Nationwide

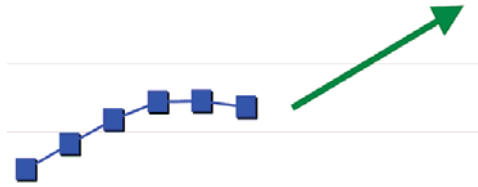


**Graph 1**

Nearly two-thirds of all listening to public radio is done by Americans with at least one college degree. Have Arbitron's estimates of their radio use gotten less reliable?

PPDV (persons per diary value) is inversely related to reliability: the lower a PPDV, the more reliable an audience estimate (all other things held equal).

There has been no decline in reliability since 2003 that would explain public radio's audience loss.



## Method, Sampling, & Response

Denial is a natural first response to bad news. So it's no surprise that the first response most broadcasters have when presented with audience loss is to deny Arbitron's ability to get it right.

They missed our listeners. They gave us a bad book.

Blaming the messenger isn't a legitimate response. But in matters of survey research, it is essential to make certain that the method is legitimate.

Arbitron employs staff specialists to deal with these questions. It's good customer service, and it's a good quality check on the product.

In most cases the analysts find nothing in the sample's distribution that violates the statistical assumption of randomness. Sometimes they find anomalies; an inordinately large (or small) number of diaries from a single zip code, for instance. And while the biasing impact is debatable, such oddities can throw a station's local estimates into question.

So despite the care with which Arbitron places its sample and balances its returns, the laws of statistics do allow a goofy local book to get out every so often.

Yet with 400,000 diaries in each Spring Nationwide – 50,000 of which

tune to at least one public station – the law of big numbers virtually ensures reliable *national* estimates ... unless, of course, some underlying, systemic shortcoming in method, sampling, or response was undermining the whole deal.

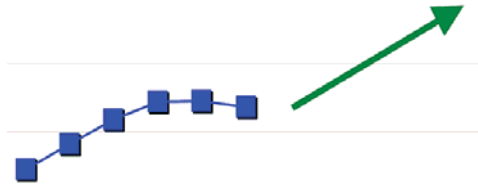
For this report, we analyzed Arbitron's **method**. Have there been changes in the diaries, or in any aspect of Arbitron's information-gathering techniques, that could have a substantive impact on our audience estimates?

We examined issues surrounding Arbitron's **sample frame**. The issue of cell-phone-only households is hot right now, and we wanted to see if that could explain anything.

Finally, we explored **consent rate** and **return rate** – the factors that determine **response rate** – the long-term decline of which is a significant concern for Arbitron and its clients. Could the resulting non-response bias account for an apparent loss of audience that may in fact not be real?

The short answer: **While not perfect, Arbitron's national numbers are sufficiently reliable. Public radio really has lost audience. The momentum driving its long run of national audience growth has indeed disappeared.**

The long answer follows.



### **The Measurement Instrument**

From time to time Arbitron improves the way it gathers or processes information about radio listening.

Such methodological adjustments have caused systemic changes in the size and nature of Arbitron's estimates in the past.

Take, for instance, Arbitron's introduction of the daypart diary in the 1980s. It was just like the diary before it, except that lines were added at daypart breaks to help people communicate what times of day they were really listening.

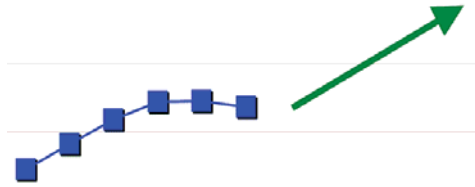
Among other things, it was expected to reduce twelve o'clock confusion: does an entry for 12:00 mean midnight or noon? The daypart diary would tell.

The change had its intended effect. Measurement got better. Questionably long stretches of listening that previously ended at midnight now ended at noon.

And as a result, reported listening levels in some dayparts declined slightly. Better measurement doesn't inevitably yield larger numbers.

Has Arbitron instituted similar improvements lately that could explain the reported loss in public radio's audience?

We raised this question with Dr. Ed Cohen, Arbitron's VP for Domestic Radio Research. He confirmed that no substantive methodological changes have occurred that would affect the syndicated estimates in this period.



### Obeying the Laws of Probability

Survey researchers strike a deal with the laws of probability when employing their craft. They don't have to administer their questionnaire to every person in the population; they can just sample a few. But in return, every member of that population must have an equal chance of being included in the survey.

All sorts of problems arise if that deal isn't fulfilled. For instance, academic studies are notorious for administering surveys on readily-available undergrads – people notoriously unrepresentative of America's population as a whole.

Another problem – non-response bias – arises when the people who don't respond to the survey are substantively different from those who do. For instance, a written survey about literacy is likely to suffer significant non-response

bias among those who cannot read the questionnaire.

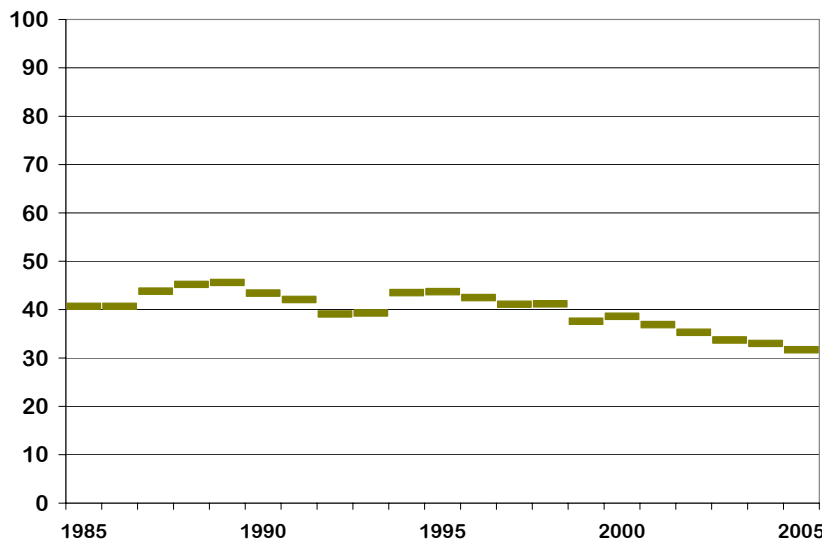
The survey would accurately report that 100% of its respondents can read. But this finding certainly can't be projected to all Americans. Such is the impact of non-response bias.

Arbitron is refreshingly open about these issues and its ongoing efforts to address them.

But this study's concern isn't about long-standing deficiencies that Arbitron itself acknowledges. Our concern is whether there have been significant changes, both recent and drastic, that might throw into doubt the reality of public radio's reported decline in audience since 2003.

We review each essential component of survey research to arrive at the answer.

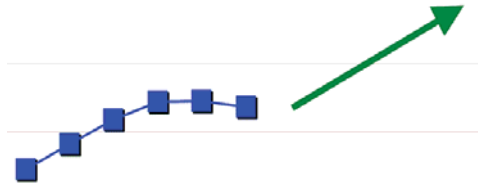
**Arbitron Response Rate**  
All Persons 12+ in the Designated Sample  
Spring Surveys



**Graph 2**

Response rate is in decline, but no more so after 2003 than before.

While not as high as it once was, Arbitron's response rate is impressive by any modern standard – especially given the increasing difficulty of telephone contact and the level of commitment it asks of diary keepers.



## Sample Frame

Arbitron's sample frame is based on households with landline telephones. It does not include group quarters (e.g. dormitories, barracks, or prisons) or cell-phone-only households.

As cell phones become ubiquitous, landlines may become less of a necessity – particularly among persons setting up new households.

Since “cell-phone-only” Americans are not in Arbitron's sample frame they are a source of increasing concern in the industry.

Arbitron's internal research into this problem leads the field. It shows nearly seven percent of all Americans (double that among 18-24 year olds) cannot be reached by residential land lines today.

It also shows that the exclusion of cell-phone-only Americans from its sample frame has a miniscule impact on audience estimates. And the marginal effects that do exist are in demos that are generally too young for public radio's programming.

Because the cell-phone-only issue remains a concern for commercial stations with younger formats, and because the cell-phone-only segment of the population is expected to grow, Arbitron has committed to incorporating a cell-phone-only sample frame into its studies by 2008.

But as the cell-phone-only population does not represent the bulk of public radio's listeners or contribute the bulk of its listening today, it is not a cause of public radio's audience woes.

No other sample frame issues have arisen since 2003 that would explain away public radio's loss of audience and growth momentum.

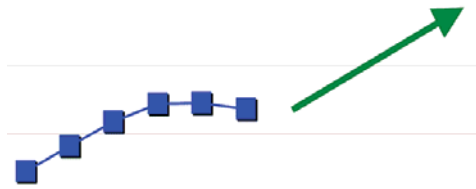
## Consent Rate

Ed Sullivan's “Talk of the Town” was television's most-viewed program in 1949 when Arbitron (then American Research Bureau) reported its first audience estimates.

Then, and throughout the company's first few decades, Americans actually appreciated phone calls asking them to participate in media surveys. Someone cared what they thought!

The advent of telemarketing and the explosion of telephone polling have trampled Americans' willingness to participate in Arbitron's surveys. So have technologies that allow people to screen their calls, such as voicemail and caller identification.

We don't know our listeners' relative willingness to participate in telephone surveys. We do know, however, that their educational attainment makes them prime candidates for screening technologies; they tend to have both the resources to purchase these services and the skills to use them effectively.

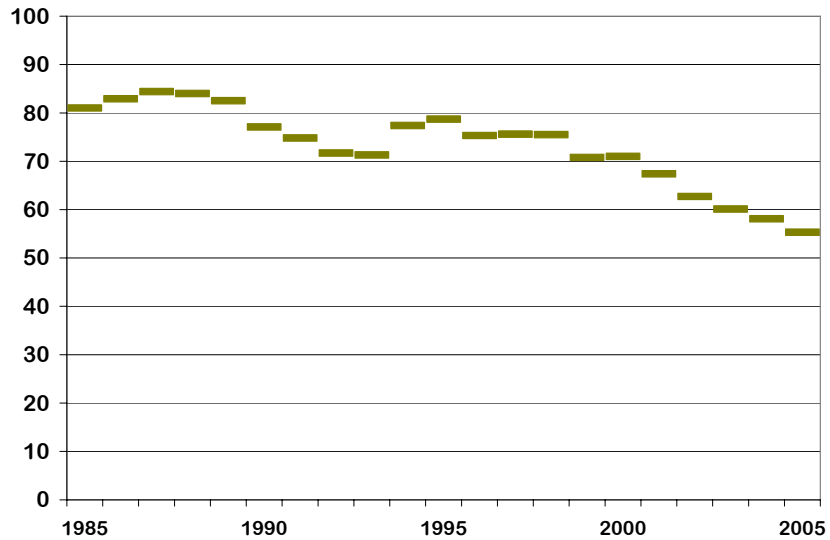


**Graph 3**

**Arbitron Consent Rate**  
All Households in the Designated Sample  
Spring Surveys

Consent rate is the percent of households in the designated sample that agree to participate in Arbitron's survey.

Consent rate for all research firms is in steady decline as Americans become more protective of their privacy and time; they are harder than ever to reach by phone, more wary of phone calls from commercial sources, and less willing to participate in surveys of any kind.



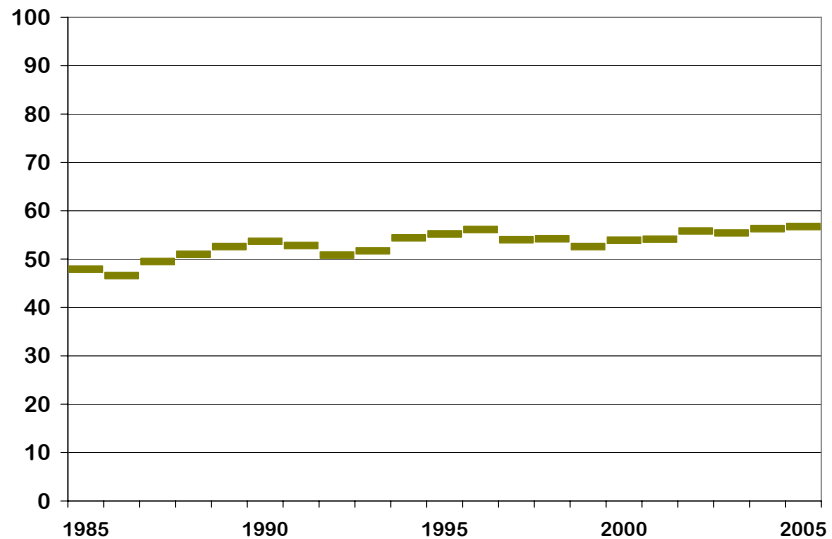
**Graph 4**

**Arbitron Return Rate**  
All Persons 12+ in the Designated Sample  
Spring Surveys

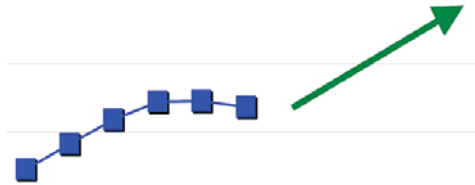
Return rate is the percent of consenters returning usable diaries to Arbitron.

Return rate is actually increasing – a direct result of Arbitron's ongoing research and experimentation.

The decline in response rate is attributable to the increasing difficulty in getting household consent in the first place.







Ideally, we'd examine the consent rate among college-educated Americans by taking the number of college-educated respondents and dividing by the number of college-educated persons in the designated sample. Unfortunately, Arbitron does not know the latter number because it can't ascertain educational attainment until a person has returned a usable diary.

It is remotely possible that the consent rate suddenly dropped for college educated households. But we see no evidence that the overall decline in consent rate has accelerated appreciably since 2003.

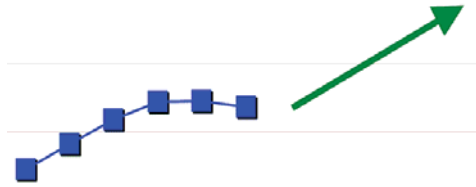
### **Return Rate**

The return rate is the percent of those agreeing to participate in the survey who submitted a usable diary to Arbitron.

Arbitron maintains an aggressive program of ongoing research and experimentation dedicated to optimizing return rates. As a result, its return rates are actually in long term ascent.

Better return rates since 2003 certainly don't support the notion that public radio's audience loss is due to sampling.

Consent Rate	the percent of sample who said they would keep an Arbitron diary
× <u>Return Rate</u>	the percent of consenters who returned a usable diary to Arbitron
Response Rate	the percent of the sample who returned a usable diary to Arbitron



## Proportionality & Weighting

Proportionality is just a 15-letter word indicating how well a sample of diary keepers represents the population. It is important to our study because it estimates something that cannot be pinned down exactly: the reliability of Arbitron's estimates for "our listeners", which we proxy as Americans with at least one college degree.

For the sake of illustration, assume that Arbitron surveys a population of 240,000 men and 240,000 women with an in-tab sample of 1000 diaries. If response were truly random, we'd expect something close to 500 diaries from men, 500 from women.

But response isn't random. We might get 600 diaries from women and 400 from men. In such a case our sample has "proportionality" issues; women are over-represented, men are under-represented.

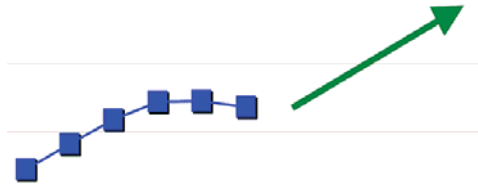
This is a real problem. Arbitron has long-standing proportionality issues with certain demographic groups – men between the ages of 18 and 24, for instance. These issues impact the reliability of the audience estimates for that demographic group.

When returns are not proportional to the population, they must be adjusted through a process called **weighting**. In the example above, the women's diaries are weighted downward by a factor of .83, or 50%/60% (their prevalence in the population divided by their prevalence in the sample). Similarly, the men's surveys are weighted upward by a factor of 1.20 (50%/40%).

When projecting these diaries to the population at large, each woman's diary will represent 400 women (240,000 women in the population divided by 600 diaries in-tab). Similarly, each man's diary will represent 600 men (240,000 divided by 400).

These values – 400 for women and 600 for men – are called **PPDVs (persons per diary values)**. As its name implies, PPDV reports the number of persons represented by each diary. It is simply the population divided by the number of diaries in the sample that represent this population.

PPDVs can serve as indicators of the relative reliability of Arbitron's estimates. In this case, the estimates for women would be generally more reliable than the estimates for men.



## Reliability

The information presented in the last several pages – pretty wonky stuff, to be sure – prepares us to address the central research questions.

Can we trust the Arbitron estimates? Is our national audience truly down? Is the momentum in our audience growth actually lost?

It all comes down to reliability – a complex construct that can be quantified only in probabilistic terms (when it can be quantified at all).

PPDVs are not reliability estimates *per se*, but they are good proxies. They are direct results of the sampling and surveying processes, and thereby embody their shortcomings and successes.

Were we to find a sharp rise in PPDVs since 2003, we could not rule out the possibility that decreased reliability of audience estimates may in fact have caused apparent rather than real audience loss.

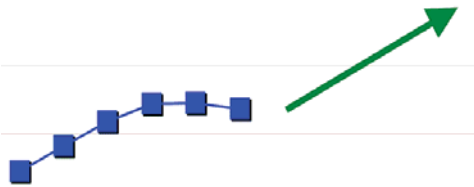
Of particular concern is the reliability of estimates for Americans with at least one college degree. Nearly two-thirds (64%) of all listening to public radio is by this segment of the population.

Here's what we find (Graphs 5 and 6 on the following page).

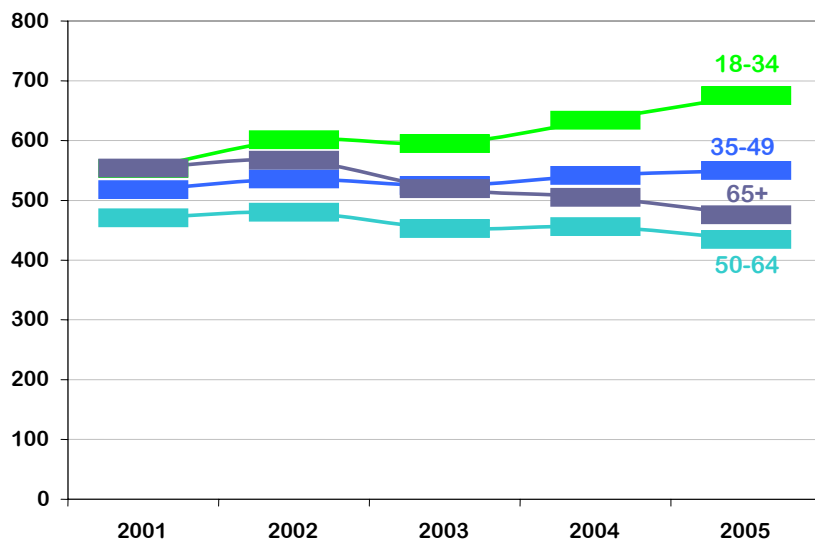
**The reliability of national audience estimates for college graduates over 35 years of age has remained essentially stable or gotten better over the last few years (lower PPDVs).**

Younger Americans are more difficult than others to get in tab. And this remains a concern among Arbitron and its commercial radio clients.

But given that 86 percent of all public radio listening is by persons over the age of 35, and given that there's been no decline in 35-plus reliability coincident with public radio's reported audience loss, we conclude that **Arbitron's national audience estimates are as reliable as they were between 2000 and 2003, when listening was still in ascent.**



**Average PPDV**  
All Radio Listeners  
Arbitron Spring Nationwide



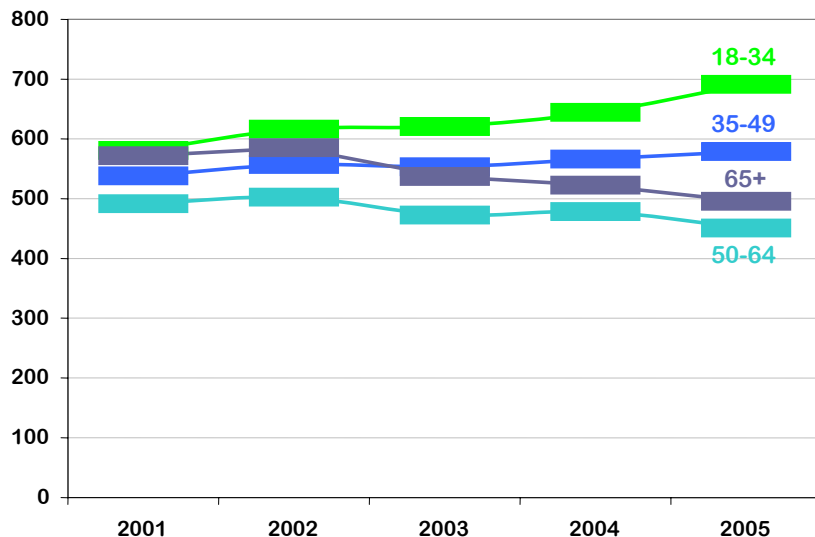
**Graphs 5 & 6**

The increasing difficulty of getting young adults into Arbitron's sample is shown on these two graphs.

So is the stability of audience estimates for persons over the age of 35, where most public radio listening is done.

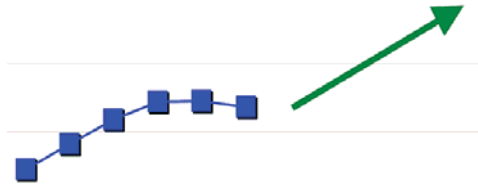
PPDV (persons per diary value) is inversely related to reliability: other things held equal, the lower a PPDV, the more reliable an audience estimate.

**Average PPDV of College Graduates**  
All Radio Listeners  
Arbitron Spring Nationwide



The top graph shows that PPDVs are holding steady among Americans 35-49 and actually improving among Americans over the age of 50.

The bottom graph shows the same patterns for Americans with at least one college degree. PPDVs are slightly higher across the board due to the increased difficulty in getting these Americans in the sample (refer to Graph 1).



## Reinvigorating Public Radio

We undertook this exercise as a professional precaution. If we had found Arbitron's data to be unreliable, there would be no reason to analyze them further.

As we were testing Arbitron's research, we concurrently reviewed several influential studies (both public and proprietary) commissioned by public broadcasters from other research and consulting firms.

We've come to a realization that we believe is relevant to public radio's quest to reinvigorate both its public service and public support.

**Arbitron is the gold standard in radio audience methodology.** Sure, the publication of its declining response and consent rates opens it up to criticism. But Arbitron's transparency in its sampling and operations – particularly its adherence to the dictates of probability sampling – only add to the trustworthiness of its reports.

**Too many other companies now selling research to public radio do not reveal the design of their sample frames, their controls against response bias, or their procedures to assure the reliability and the integrity of their final products.**

The dictates of the sample frame are the first thing taught in any survey course. Yet many studies now on the desks of decision-makers are based on non-pro-

jectable convenience or volunteer samples (such as pop-up website surveys). Yes, these samples are cheap – but they are neither projectable nor adequate for making decisions on expensive and critical investments.

We find myriad other deficiencies in the variety of reports examined. In far too many, adherence to the objective requirements taught in Surveying 101 is obviously and sadly lacking.

**Decision-makers often excuse research as “good enough” for the task at hand. A question they're apparently not asking is: when does it become “bad enough”?**

**Public broadcasters should demand full disclosure of methodology, especially with reference to sampling and response bias, for any survey or report on which they are to make crucial decisions.**

Compare the research reports on your desk to Arbitron's standards. You may find that you are being misguided by numbers that are “bad enough” by all objective, scientific criteria.

Ask yourself: “Would I make the same decisions if I knew this information was wrong?”

**It's better to have no information than bad information. The magnitude of public radio's current challenges requires that the research on which it relies rise to significantly higher standards.**