# 17

# Internet and mobile ratings panels[1]

**Philip M. Napoli[1], Mario Callegaro[2], and Paul J. Lavrakas[3]**

[1]*Fordham University, USA*

[2]*Google, UK*

[3]*Independent consultant*

## 17.1 Introduction

Internet ratings panels represent a distinctive type of online panel. Internet ratings panels are designed to provide systematic data on audience size, demographics, and exposure patterns for individual web sites and mobile applications. These ratings data play a central role in the transactions that take place between online content providers and advertisers, in addition to serving a variety of other analytical purposes.

These panels are typically constructed and managed by third-party measurement firms. These firms provide data on a subscription basis to web site operators, advertisers, media buyers, academic researchers, policy analysts, and any other individuals or organizations with an interest in understanding the size, demographic composition, and behavioral patterns of online audiences.

Ratings panels are used to provide key web site performance criteria such as the number of unique audience members visiting a web site in a given month; the active reach (which is the percentage of web users who visited a site in a given month); and the time per person

[1] We thank Sherrill Mane from the Interactive Advertising Bureau (IAB) for her support in writing this chapter. Alexei Zverovitch (Google) helped with the Java code and parsing of the XML output from Alexa Top Sites in order to create Figure 17.1.

*Online Panel Research: A Data Quality Perspective*, First Edition.
Edited by Mario Callegaro, Reg Baker, Jelke Bethlehem, Anja S. Göritz, Jon A. Krosnick and Paul J. Lavrakas.

spent on the site (see Nielsen, 2012b). Such performance metrics can also be segmented via demographic categories.

This chapter examines how Internet (PC and mobile) ratings panels are constructed, managed, and utilized. We provide an overview of the history and evolution of Internet/mobile ratings panels. It examines the methodological challenges associated with creating and maintaining accurate and reliable Internet/mobile ratings panels. Research that has assessed the accuracy and validity of online panel data is critically examined as well as research that illustrates the type of scholarly and applied research questions that can be investigated using online ratings panel data. The chapter concludes with a discussion of the future of online ratings panels within the rapidly evolving field of Internet audience measurement.

## 17.2 History and development of Internet ratings panels

The commercialization of the Internet – and (equally important) the development of the World Wide Web – began in the early to mid-1990s. The first banner ad appeared on the web site of *HotWired* magazine in 1995 (Bermejo, 2009). It is around this time that the first systems for providing audience ratings for the fledgling medium emerged (see Flynn, 1995). Many of these systems relied on online panels to produce their audience ratings (see Coffey, 2001).

However, even in the earliest days of the web, online panels were just one of a variety of possible ways of producing Internet audience ratings. Other approaches, such as relying on the analysis of web site server logs, or analyzing web traffic data aggregated by ISPs, have also been – and continue to be – employed (see, e.g., Bennett, et al., 2009; Bermejo, 2007).

The earliest online ratings panels began in 1995 with a service called PC-Meter. Created by the NPD Group, the PC-Meter service operated from a 500-household panel (Berjmejo, 2007). This earliest panel highlights one of the persistent challenges in the provision of Internet audience ratings – creating and maintaining a representative panel of sufficient size to account for the tremendous number of web sites vying for audience attention. Even in the early days of the web in the mid-1990s, a 500-household panel would be capable of providing meaningful audience size and demographic estimates for a very small fraction of the web sites in operation.

By 1997, additional ratings providers had entered the marketplace – Relevant Knowledge and NetRatings; followed in 1998 by NetValue, and in 1999 by a measurement service provided by computer sales and marketing firm PC Data (Bermejo, 2007). As often happens in the field of audience ratings, consolidation took place fairly quickly (Napoli, 2003). Many of the early providers of panel-based audience ratings data quickly merged; and, perhaps most significant, television ratings giant Nielsen moved into the field, partnering with NetRatings in 1998 (Bermejo, 2007). Nielsen would later acquire the firm, with Nielsen NetRatings continuing to serve as one of the leading providers of Internet audience ratings today (now under the name Nielsen Online), with the panel-centric measurement service now operating under the name Nielsen Netview.

As might be expected in the nascent stages of a new field of audience measurement, these early online ratings providers employed a number of different methodological approaches in the construction and management of their panels. Early Internet audience measurement firm Media Metrix, for instance, employed an approach involving installing software in the operating systems of the computers of participating panelists (Bermejo, 2007). NetValue, in contrast, employed an approach that involved installing measurement software at the level of the Internet protocol suite (TCP/IP). And finally, the measurement system

developed and employed by Nielsen NetRatings (now Nielsen Online) was installed at the browser level.

The implications of these different approaches to installing the measurement software were significant. Software installed in the operating system, for instance, could gather information on all computer activities (including non-Internet-related activities), but could not detect, for instance, the distinct elements within individual web pages (such as banner ads). Measurement software installed at the TCP/IP level could monitor any form of online communication (including instant messaging, email, etc.), and could detect all of the distinct elements within individual web pages. However, this approach could not measure the time spent on the last page that was viewed in the browser, since the end of time spent on a web page was determined via the next page request (Bermejo, 2007). Measurement software installed at the browser level was also capable of gathering information on individual web page elements (such as banner ads), but could not gather information about other forms of online activity (such as instant messaging, emailing, etc.). This approach also was incapable of measuring the duration of time spent on the last page requested.

What ultimately emerged from this competitive environment were two firms that, for some time, championed somewhat different (though now converging) approaches to the construction and management of online ratings panels, and the production of Internet audience ratings estimates. Nielsen NetRatings began with (and until recently focused exclusively on) a system based upon a traditional random sample of Internet users, recruited via traditional random digit dialing recruitment procedures (see Goosey, 2003), albeit with extremely low response rates. This panel includes both home- and work-based participants, and has peaked in size at around 30000 individuals (Bennett, Douglas, Rogers, & Broussard, 2009).

In contrast to Nielsen and its traditional sampling approach, leading competitor comScore pioneered "Global Macro-Panels" (Bermejo, 2007, p. 161). By 2007, comScore had managed an online panel of over 1.5 million Internet users (Bermejo, 2007). Today, that number exceeds two million (comScore, 2012), though gained with extremely low participation rates. Moreover, smartphone users are now added to the panel together with tablet owners and game consoles and Smart TVs (comScore, 2013). Panelists are recruited online, via banner ads placed on web sites that collaborate with comScore. Incentives for participation include faster browsing speeds, antivirus protection, participation in a sweepstakes, and even offers by comScore to plant trees in exchange for participation (Abraham, 2012; Fulgoni, 2003). Nielsen has since adopted a similar approach, employing a larger non-random sample alongside its traditional random sample (Bennett et al., 2009; Nielsen, 2012a).

It is important to emphasize that this sampling approach, involving very large, non-random samples, has proved controversial. Participants are not always clear that they are agreeing to have their online behaviors measured. And in some instances measurement firms have gathered additional user information (passwords, financial data) without the participants' consent. Occurrences such as these have created a somewhat blurred line between online audience measurement and spyware (see, e.g., Levine & Finkle, 2011).

Obviously, such a recruitment approach lacks the traditional sampling rigor associated with panel-based research. Data from the smaller panel, along with RDD enumeration surveys are employed to weight the data obtained from the larger sample. This strategy of employing various types of "calibration" panels to weight data obtained by other means that may either lack the ability to be representative, or may simply lack the ability to obtain demographic information about the audience, is now somewhat standard.

comScore, for instance, like Nielsen, maintains a smaller, 120000-person "calibration" panel, recruited via traditional random digit dialing methods. This panel is comprised of

50000 home-based participants, 50000 work-based participants, and 200000 university-based participants (Bennett et al., 2009). This panel is used to weight the results obtained from the larger panel of participants recruited online (Cook & Pettit, 2009).

A more recent entrant into the Internet audience measurement arena, Compete, maintains a panel with a sample size of roughly two million participants. Compete engages in what has been described as "panelist multisourcing," which involves gathering data from the company's own proprietary panels, which is combined with licensed clickstream data obtained from third party partners such as ISPs (Bennett et al., 2009, p. 169). Via this approach, Compete (2011) claims to be able to provide detailed audience information for approximately one million web sites (see Bennett et al., 2009).

Some tech companies such as Facebook and Google have started building panels in order to be able to meet their advertising clients' needs. Facebook partnered with Nielsen to measure their effectiveness of internet advertising (Nielsen, 2009) and also more recently with Datalogix (Reynolds, 2012) to measure offline advertising. In the Facebook case, the pool of Facebook users is considered their "panel." In other words, the concept of a panel is qualitatively different than what else has been discussed in this chapter. Google is experimenting with online panel partners such as Gfk-Knowledge Networks in the United States (Arini, 2012b; (http://www.google.com/landing/screenwisepanel/) and with Kantar in the United Kingdom (Connected Vision 2011) in order to provide advertisers with an additional way to measure online advertising. For example, a panel was launched in conjunction with the 2012 Olympics and in collaboration with the NBC network in the United States. This panel of approximately 3000 households measured Olympics media consumption from PCs, tablets, TVs and smartphones (Spangler, 2012; Steinberg, 2012).

Finally, it is important to emphasize that the Internet audience has been, in its relatively short history, something of a moving target, as technological, demographic, and behavioral changes have consistently posed new challenges to effectively capturing, aggregating, and reporting audience exposure to online content options. This is why, for instance, panel-based measurement services must continually engage in enumeration surveys in order to try to maintain accurate and up-to-date information on the size and demographic characteristics of the universe of Internet users (something that changes from month to month as Internet penetration across various access technologies continues to advance; see Cook & Pettit, 2009).

## 17.3 Recruitment and panel cooperation

By recruitment, we mean what happens during the initial stages of creating an audience measurement panel when the online panel researchers try to gain *cooperation* from potential panel members by getting them to "join" the panel. During this stage the panel member (1) must be found to be eligible to join the panel; (2) must agree to do so, including agreeing to all privacy and confidentiality policies; and (3) must engage in the set-up (aka "installation") behaviors that allow for the start of the measurement of her/his Internet usage. Exactly what "join" and "cooperation" and "installation" mean vary by panel and are discussed below.

Of note, we differentiate the recruitment (cooperation-gaining) stage from the notion of maintaining *compliance* among panel members, which refers to the on-going behavior of those who have qualified to participate, and who have agreed to cooperate with what is asked of them, and who actually have started to do so. Longer-term compliance in the panel often presents severe and separate problems for the panel vendor. These are addressed in the next section on panel attrition.

Recruitment of the general public into an audience measurement panel is a very challenging endeavor for several major reasons:

- the time burden on potential panel members to go through the "eligibility screener" process;
- the considerable cognitive burden on potential panel members to read, understand, and make the decision about whether to agree to what can be very convoluted and lengthy confidentiality/privacy agreements the panel vendor has in place,
- the time, intellectual, and/or physical burden on the panel members to install (or to have installed) the measurement system onto their computers and/or smartphones;
- the time and cognitive burden on panel members to use the system properly;
- the invasion of privacy burden on a panel member to allow a research company to learn a great detail about their Internet usage.

Because of these factors, and despite the considerable efforts which some audience measurement panels companies engage in to motivate cooperation, only a small portion of the public is willing to participate in such a panel – usually less than 5%, and quite often less than 1%. Such participation rates call into question the representativeness of any of these unweighted audience measurement panels.[2]

Recruitment sampling for Internet measurement panels occurs in one of three basic ways:

1. A new panel may be built from scratch using a probability sampling method.
2. A new panel may be built from scratch using a nonprobability sampling method.
3. A new panel may be built from members of a pre-existing panel.

### 17.3.1 Probability sampling for building a new online internet measurement panel

When a probability sampling approach is used to start a new online audience measurement panel, it often signals that the representativeness of the unweighted panel is a key concern for the panel researchers.

Currently, the two primary probability sampling designs used for recruiting these online panels are: (1) a form of dual frame (both landline and mobile/cell) random-digit dialing (RDD) sampling with telephone as the mode of contact (AAPOR, 2010); or (2) a form of multistage area sampling using an address-based frame with mail as the first mode of contact (Iannacchione, 2011). In the case of the former sampling technique, RDD surveys in the United States have undergone considerable modification during the past decade due to the considerable growth in the proportion of the population that can only be reached via mobile phone numbers; estimated at approximately 40% as of the end of 2012 (Blumberg & Luke, 2013). Thus, when an RDD sampling design is used, mobile numbers must make up a significant proportion of those who are in the final recruited sample. In the case of address-based

[2] These estimates are drawn from the extensive professional experience of one of the authors. The reader may be surprised to learn that these serious scientific concerns do not appear to harm the business prospects of most panel vendors. That is because most clients appear not to understand or care about the threats these problems cause for the reliability and validity of the data they are purchasing.

probability sampling, some research companies supplement the use of mail for the recruitment mode with telephone contact to those nonresponding addresses that can be matched to a phone number. To our knowledge, no panel vendor is as yet building an online audience measurement panel by relying on in-person recruitment as part of a probability sampling design.

### 17.3.2 Nonprobability sampling for a new online Internet measurement panel

When a nonprobability sampling approach is used to start a new online audience measurement panel, it often signals that the representativeness of the unweighted panel is not of paramount concern for the researchers. Rather, the paramount concern often seems to be to get as many "bodies" (almost regardless of who they are) to join the panel and to keep the costs of doing so as low as possible. Researchers appear to do this because of their apparent belief that the lack of representation in their unweighted panel can be adequately corrected by various statistical adjustments (or so many tell their clients).

Some of the nonprobability sampling approaches use systematic methods (e.g., filling predetermined demographic quotas) to try to recruit members into the panel, but most use various uncontrolled "sampling" methods that essentially amount to no more than what is correctly described as a haphazard self-selected convenience sample. The methods used to make potential panel members aware of the opportunity to join the new panel vary but are most often made up of Internet-placed "invitations" to do so. Thus whoever happens to notice such invitations has a chance of actually reading it and if they do, then may decide to look into what s/he is being invited to do. Recruitment invitations also may come in the form of an email or via traditional mail. In some cases, a form of telephone recruitment may also be used, but this is more likely to utilize nonrandom lists of telephone numbers rather than a technique like RDD to generate representative telephone numbers. These nonprobability sampling methods most often achieve fewer than 1% of those exposed to the invitations agreeing to join the new panel.[3]

### 17.3.3 Creating a new panel from an existing Internet measurement panel

For online panel companies that already have general population online panels in place, it is very cost effective to use their existing panel members as the sampling frame to create a new panel. Sampling from this frame can be done in numerous ways, but in almost all cases it would be foolish not to use a probability sampling design, in particular one that stratifies selection along whatever variables are judged to be key to what will be measured (e.g., geography, age, gender, race, Hispanic ethnicity, presence of children in the household, household size, etc.).

In the case of some audience measurement companies (e.g., Arbitron and Nielsen), they already have scientifically sampled panels to measure the audience for media other than the Internet. For them, it is a simple sampling task to try convert some or all of these households so that they agree to have their online behavior measured in addition to what other media exposure/consumption is measured. These "cross-media" or multimedia panels have the great advantage of measuring media behaviors on more than one medium from the same people. Their disadvantages are that it is not a random subset of existing panel members who will agree to allow their Internet behaviors to be measured in addition to the media behavior they

[3] This estimate comes from the extensive professional experience of one of the authors. Such response data are typically kept proprietary by online audience measurement firms.

previously agreed to have measured. But, as a counter to this disadvantage, these companies are in the excellent position to study the presence of nonresponse bias (its size and nature) in the new audience measurement panel because they already know so much about each panel member/household in the entire frame from which the new panel is being sampled.

When these existing panels have been formed via probability sampling methods, the original response rates can be relatively high (e.g., in the 20%–40% range for Nielsen's People Meter panels). But when panel households are asked to also allow their Internet usage to be measured, the subset of exiting panel members who agree to cooperate is not a random subset of the original panel sample and typically is well less than 50% of those who are asked to do so.[4]

### 17.3.4 Screening for eligibility, privacy and confidentiality agreements, gaining cooperation, and installing the measurement system

Regardless of which sampling design is used to get people into an Internet measurement panel, the first step in the actual recruitment of these individuals (and possibly the recruitment of others in their households) is to determine whether they meet the panel's eligibility requirements. Such requirements typically are technology-based, but other than being a resident of the geo-political area being measured and age, they are not often demographic-based.

To accomplish this, a person is asked to complete an "intake" questionnaire that first determines one's eligibility. For those who are found to be eligible, they typically are asked to read and agree to whatever privacy requirements and confidentiality pledges the panel vendor (and possibly its client) has in place. It is at this point where a number of people refuse to continue. This is a difficult tradeoff for Internet measurement companies, because they need and want to protect themselves from various privacy-related liabilities and yet by disclosing in great detail all the data that will be gathered and what they will do with the data, they scare away a good proportion of people from agreeing to join the panel. And there is no good reason to assume that the people who drop out of the recruitment process at this point are a random subset of those who began the process.

For those who agree to the privacy/confidentiality stipulations, they then are informed in more detail about what "joining" the panel entails and then are asked to agree to do so. The burden of actually serving in such a panel typically is not an onerous one, and in some cases is essentially a one-time shot with little or no continued effort required from the panel member.

After they agree to join, there are two basic means by which Internet usage is measured. One system is entirely software-based, in which an app or other type of software is downloaded or otherwise installed onto their computers, tablets, and/or smartphones. The other system involves the installation and use of a piece of hardware in the household – an Internet "meter" – such as a wireless router, in addition to software that may need to be installed onto the devices someone uses to reach the Internet. For the most part, the operation of these systems is invisible to the panel member, with the occasional need to update demographic data and answer some topic-specific questionnaires, all of which is done online.

### 17.3.5 Motivating cooperation

As noted above, there are a great many hurdles to overcome between initially contacting a potential panel member and gaining a final commitment for cooperation from that person.

[4] These data come from the extensive professional experience of one of the authors. As was noted in the previous footnote, such response rates are typically kept confidential by online audience measurement firms.

Lavrakas and colleagues (2012) have advised that the entire process is best viewed as a multi-step challenge and that recruitment procedures should be devised and implemented for achieving success at each of the steps. Thus, the notion of addressing only the end-goal – gaining final cooperation – is not prudent.

Instead the strategies that are used to eventually gain cooperation are dependent on successfully advancing a potential panel member along the entire continuum from sampling to usage of the Internet measurement system. So, for example, offering a contingent incentive only to those persons who are found to be qualified and eventually agree to cooperate is not enough. Ideally sampled individuals will be motivated to accomplish each step and then move onto and through the next step: starting with visiting the web site to begin the intake questionnaire, completing the eligibility section, agreeing to the confidentiality/privacy stipulations, agreeing to join the panel, installing the measurement systems, and then starting to use them. Experience shows that achieving cooperation will depend directly on:

- how interested the person is in what is being measured;
- how important s/he thinks the measurement of this behavior is and who s/he perceives will benefit, and by how much, from the existence of the panel;
- how much s/he trusts that no harm will come will come to her/him or their household from participating;
- how much initial and on-going burden s/he will experience to be a panel member;
- the "value" of the material and nonmaterial incentives s/he and her/his household will receive for joining and staying active in the panel.

Internet usage measurement companies that strive to get only "a lot of bodies" into their panels will use the least costly and therefore the least effective techniques for recruitment, thinking that the cooperation rate does not matter. But the science of building such online audience panels is still in its infancy and no one knows for certain whether it is more cost-effective to build a panel by spending more or less money on each successfully recruited member.

## 17.4 Compliance and panel attrition

As noted above, compliance is the term that defines how well Internet audience measurement panel members "behave" after joining the panel, including how long they remain as members compared to the length of time for which they originally agreed to stay in the panel. The latter speaks to the issue of *panel attrition*, i.e., whether, and why, someone drops out or is forced out before their original commitment has ended.

From an on-going compliance standpoint, experience shows that within a general population panel, there will be three groups of panelists: (1) those who comply well with what they agreed to do and need little or no encouragement from the panel managers; (2) those who comply well enough, but need occasional encouragement (intervention) from the panel managers, and (3) those who do not comply well at all and need a lot of intervention by the panel managers.

In the extreme case, some among this last group will have such high noncompliance and will require so much attention from the panel managers that a point will be reached where they are likely to be forced to drop out of the panel. This process has been euphemistically

called by some "unforced attrition," because it happens before the panelists reach the point (if there is one in a particular panel) when they would be forced to leave the panel due to their tenure reaching the maximum time permitted for the panel.

There are two broad types on noncompliance – often termed *faulting* – within audience measurement panels. One form of faulting is associated with technical problems with the system that is used to measure a panelist's Internet usage. These technical faults may occur because the system hardware and/or software itself fails to work reliably. In these cases, the system needs to be repaired or new hardware/software needs to be installed to bring the panelist back into compliance (i.e., eliminate the faulting). However, in some instances these technical faults occur because of some behavior in which the panelist has or has not engaged. Examples include problems such as buying a new computer or smartphone and not informing the panel company so that the measurement system can be properly installed on these new devices; or moving a desktop computer or the router to another room and not reconfiguring it to work properly with the measurement system. In these cases it is not the equipment's "fault" that the panelist is no longer in compliance, but rather the panel member's fault. Although these latter types of technical faults may occur with just about any panelist, experience shows that they occur much more frequently with certain types of panelists (see below).

A second category of noncompliance occurs because of behavioral problems associated with the panelist. These social or behavioral faults are associated with panelists who simply forget or refuse to continue doing what they originally agreed to do as part of their panel membership agreement. For example, some online Internet usage measurement systems may require that a panelist login to the measurement software to indicate which household member is presently using the computer. In some measurement systems the user may be prompted occasionally to re-login, in case the user has changed since the last login and the previous user failed to properly log out. In the case of smartphone software, the login–logout problem is moot as most panel measurement companies tend to assume there is only one user for a given smartphone. For some people, this requirement to login/logout is an onerous burden, and thus they will be noncompliant for various times when they are not properly logged in as themselves.

A related, but different type of behavioral faulting occurs in some measurement systems in which Internet users can log themselves in as a "guest" without disclosing anything else about themselves. Were legitimate panel members to do this because they did not want the Internet usage in which they were about to engage to be directly associated with them, they would technically be noncompliant for that time of Internet usage. Were they to do this over and over again selectively (and there is good reason to believe that some people do this quite often, e.g., those panel members who do not want anyone else to know every time they visit porn sites), then this would constitute a social/behavioral form of noncompliance.

Unfortunately for audience measurement companies and their clients, it is not a random subgroup of the panel membership that consistently has technical and/or behavioral problems with noncompliance (Holden, Heng, Lavrakas, Bell, & Flizik, 2006). Experience has shown that low compliance in audience measurement panels is likely to be correlated with panel members who are younger (< 35 years), African Americans, Spanish-Dominant Hispanics, have lower income, have lower educational attainment, have at least one child (< 18 years) living in their household, and/or are members of large-sized households. To the extent this happens, the data being gathered each day about the panel's Internet usage will be biased, as there will be missing data that are "not missing at random."

To address these issues, Internet audience measurement companies can implement various strategies to try to reduce or possibly eliminate a panel member's faulting. These efforts can be as simple as sending the noncompliant person/household a reminder email. At the other end of the spectrum is the deployment of a "personal coaching" program, whereby a member of the panel management staff telephones or even may visit in person the problem panelists and attempt to motivate them to improve compliance (Holden et al., 2006). Depending on the approaches used, this can be an extremely expensive enterprise. To our knowledge, very few business models for Internet audience measurement companies support doing more than sending email reminders alerting panelists of their noncompliance problems, if even that. Instead, the *modus operandi* for most of these panels for reasons of costs seems to be to simply drop those panelists who are consistently noncompliant, and replace their number, if not their "type," with new panelists.

Another form of attrition is linked to those who drop out of an online Internet measurement panel prematurely (before their originally agreed to tenure ends) because of a growing concern about the invasion of privacy to which they are subjecting themselves by allowing their Internet behavior to be measured. These are likely to be individuals who originally joined the panel with some reluctance because of their privacy concerns, and these concerns begin to "eat away at them" over time. They finally reach a point where they think "enough is enough," even if they have no tangible evidence that they have been harmed by their participation in the panel. There is no valid reason to believe that those who leave Internet measurement panels for these reasons are a random subset of the entire panel membership, though we are unaware of any definitive empirical evidence on this issue.

## 17.5 Measurement issues

There are many measurement issues to consider when measuring audience behavior with online panels (Media Rating Council, 2007). Measurement issues can be classified into the following three broad groups: (1) coverage of Internet access points; (2) confounding who is measured; and (3) differences in metering technologies.

### 17.5.1 Coverage of Internet access points

With the proliferation of devices that can access the Internet (tablets, smartphones, laptop and desktop computers at both home and work), the challenge is to measure all of them. For example, Internet browsing at work is not easily captured and linked back to the same household panel member. Tablets are becoming more and more prevalent in households and many studies are converging on the finding that smartphones and tablets are used as a second screen when watching TV (Smith & Boyles, 2012). In the United States, for example, half of the adult mobile phone owners have used their phone for engagement, diversion (e.g., during commercial breaks), and fact checking when watching TV (Smith & Boyles, 2012). In another report (Nielsen, 2011), it was estimated that "36% of people 35–54 and 44% of people 55–64 use their tablets to dive deeper into the TV program they are currently watching" (p. 4). These data points are important to understanding the need to measure every device in order to have a complete picture of Internet behaviors.

According to recent estimates, nearly half of the US adult population owns a smartphone, while 25% own a tablet and 19% an e-book reader (Raine, 2012; Raine & Duggan, 2012). In the United Kingdom, for example, 54% of adults 16 and older own a smartphone while tablet

ownership is at 16% at a household level (Ofcom, 2013). These technologies are becoming an increasingly prominent means by which individuals go online. And, in some countries, these devices are far and away the *primary* means by which the bulk of the population goes online (Napoli & Obar, in press).

According to current assessments, however, the measurement of online audiences via mobile platforms has not kept pace with the rate at which audiences are using these devices to go online (see, e.g., Gluck, 2011). For instance, measurement in this area has yet to fully move to the third party provider model that characterizes other media, and that is generally considered essential for generating the necessary levels of advertiser trust and confidence in the audience estimates. Further, existing measurement systems remain, at this point, largely unable to describe the mobile audience in terms of unique visitors (Gluck, 2011), which, of course, undermines the audience targeting models traditionally employed by advertisers, as well as the ability to conduct comparative analyses across media platforms. As research by Gluck (2011) has indicated, "Publishers and advertisers feel frustrated that mobile, to date, does not offer the same metrics and accountability found in PC-based digital media" (p. 6).

This situation is due, in part, to the greater complexity associated with measuring audience behavior via mobile devices. This complexity arises from a range of factors, including substantial variability in terms of the devices, operating systems, and carriers that need to be accounted for in the measurement process; as well as more complex privacy issues that arise within the context of mobile-based Internet audience measurement; and the prominence of mobile applications (i.e., apps) rather than web pages as the mechanism via which individuals engage with online content (Gluck, 2011).

Panel-based measurement of Internet audiences via mobile devices is very much in its infancy. Nielsen, for instance, released its first mobile rankings of web sites and mobile applications in late 2011 (see BusinessWire, 2011). These rankings were derived in part from a panel of 5000 smartphones installed with metering software (BusinessWire, 2011). comScore launched a panel-based mobile measurement system, Mobile Metrix 2.0, in mid-2012 (Opidee, 2012). Through these panels, data are gathered on a wide range of mobile-based activities, including application usage, web site visits, geographic location, and even data and duration of phone calls (Nielsen, 2012c). US radio ratings firm Arbitron recently launched its Mobile Trend Panel in the United States, the United Kingdom, Germany, France, Denmark and Finland. It is a consumer opt-in panel in which panel members install an app on their smartphones and tablets that tracks their Internet and phone behavior. A questionnaire can also be triggered from the app based on context-specific situations (Arbitron 2012a, 2012b).

### 17.5.2 Confounding who is measured

Internet rating panels rely on panel members and on technology to identify who is using what device. In industry language this is called "Attribution of user sessions to specific users" (Cook & Pettit, 2009). This is the case of shared devices such as laptops, desktops, and tablets. For example, Nielsen (2011) estimated that 43% of tablet users share their device with household members. One way to assess who is using what device is to require panel members to identify themselves on the device used at certain intervals. This strategy is prone to compliance issues (see above). Another method relies on technology to identify the device user. For example, comScore (Cook & Pettit, 2009) uses a patented procedure called User Demographic Reporting (UDR) that analyzes mouse and keystroke patterns in order to build a profile for each panel members that is then cross-checked with a self-identification survey.

When talking about mobile rating panels, the measurement issues are slightly different. First, there is very little sharing of smartphones among family members. Measuring Internet behaviors on smartphone generally entails panel members installing an app that tracks their Internet traffic. As with Internet panel members, compliance is an issue. Another way is to measure smartphone Internet behaviors only at home when the panel members use their smartphone connected to their home Wi-Fi and the Wi-Fi router sends data to the panel company. This methodology has the advantage of not requiring panel members to install an app and the disadvantage of not measuring smartphone Internet behavior outside the home (for a discussion of additional measurement issues, see Gluck, 2011).

## 17.6 Long tail and panel size

Perhaps the single greatest challenge to effectively utilizing panels to calculate online audience ratings involves the construction and maintenance of sufficiently large panels to account for the unprecedented levels of media and audience fragmentation that characterize the online space. Audiences are fragmented across the array of web sites available to them (see FAST/ARF, 2000; Fulgoni, 2003; Napoli, 2011a), which creates the persistent problem of effectively capturing the distribution of audience attention across these content options. The pattern that we typically see, in terms of the distribution of audience attention, is one in which a large proportion of the audience is clustered around relatively few popular sites, with the remaining audience attention very widely dispersed across an incredibly large array of relatively unpopular content options. This pattern has been famously termed the "long tail" by media analyst Chris Anderson (2008). This pattern is represented in Figure 17.1 with Alexa data as of June 2013 for the US traffic.

In Figure 17.1, we can see that to be among the top 1000 web sites, very few sites capture more than 5% of overall traffic in the United States, very quickly the majority of sites capture 1% or fewer visitors per month; and perhaps more importantly, once outside the top 100 web sites, we are well into the Internet's exceptionally long tail. For example 0.005% of traffic is captured by site rank number 877. This tail is so long that, according to one recent estimate, companies such as Nielsen and comScore can "provide detailed audience measurement for [only] fifteen to thirty thousand of the estimated 187 million available websites" (Napoli, 2011a, p. 74). The audiences for the remaining 186 million plus web sites remain essentially unmeasurable via purely panel-based measurement systems. Granted, the audiences for each of these individual web sites are incredibly small; however, in the aggregate they represent a substantial portion of online audience attention. This dynamic creates a powerful incentive for online audience measurement systems to try to reach as far down the long tail as is practically possible.

The unprecedented fragmentation of the online space means, however, that it is for all practical purposes impossible for panel-based ratings systems to provide accurate and reliable measurements of even a fraction of the totality of available web sites. The costs associated with recruiting and maintaining panels impose significant limits on the extent to which panels are able to provide detailed audience data on the web sites that populate the long tail of audience attention (Fulgoni, 2003).

Other measurement approaches, such as the analysis of server logs, offer ways around the long tail problem (Bennett et al., 2009). However, these analytical approaches are beset with their own distinctive methodological shortcomings. In particular, these measurement approaches lack the ability to accurately and reliably provide the kind of

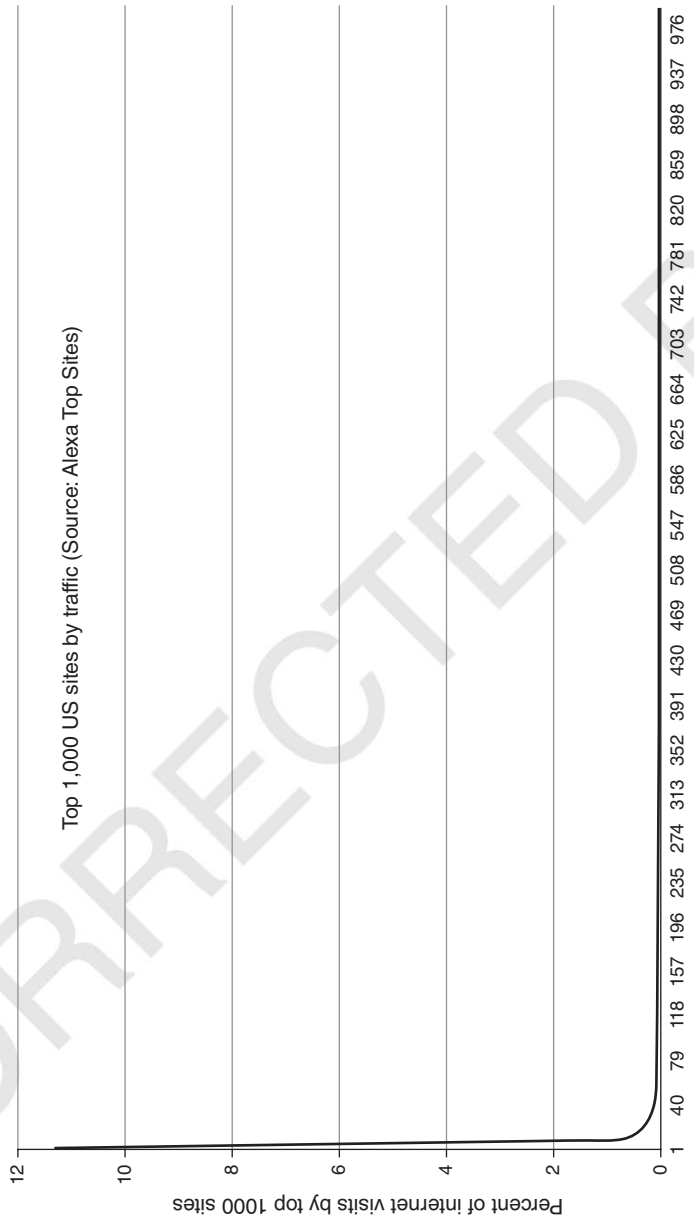


**Figure 17.1** June 2013 percentage of US traffic for the top 1,000 U.S. web sites. Reproduced by permission of of Amazon Web Services*Source:* Data from Alexa Top Sites http://aws.amazon.com/alexatopsites/.

audience demographic information that advertisers have come to demand. As the Interactive Advertising Bureau (2009) notes in their Audience Reach Measurement Guidelines,

> While it is possible for census-based measures to produce counts of Unique Users . . . the threshold of measurement difficulty for achieving this measure in a census-based environment is quite high (generally because of the difficulty of being able to identify a cookie as a unique person persistently during the measurement period).
>
> (p. 4)

Census-based approaches are used for measuring smartphone traffic as well. For example, comScore in collaboration with some UK mobile operators carriers collects anonymized census-level data traffic (comScore, n.d.).

## 17.7 Accuracy and validation studies

The history of online audience ratings has, from its very beginnings, been characterized by persistent evidence of substantial discrepancies across different measurement services (see, e.g., Fitzgerald, 2004). Anecdotal examples of extreme discrepancies in audience estimates across measurement services – even among those employing methodologically similar approaches (e.g., across panel-based measurement services) are commonplace (Story, 2007; Graves, Kelly, & Gluck, 2010). It is not uncommon, for instance, for the two leading panel-centric online ratings firms to fail to even be in agreement as to what is the second most popular online destination in a given month (Graves et al., 2010). Some analyses have contended that these discrepancies, and the uncertainty that they engender, have been the primary impediment to the growth of online advertising (see, e.g., Graves et al., 2010; Story, 2007).

Detailed audits of the accuracy and reliability of audience measurement systems are seldom made publicly available in the United States, where such audits typically are handled on a confidential basis by the Media Rating Council (MRC). While the MRC will make public the outcome of such an audit, in terms of whether the audit resulted in the MRC accrediting the measurement service, the underlying analyses typically are kept confidential. As of this writing, the MRC is engaged in audits of the leading panel-centric Internet audience measurement systems in the United States. Such audits can take into consideration each of the various stages in the process of generating audience ratings, from the enumeration survey, to the metering process, to sample recruitment and weighting (see Haering, 2002, for an example of an audit conducted on three panel-based online audience ratings services operating in France at the time).

In 2009, the Advertising Research Foundation released a fairly detailed assessment of the comScore methodology (Cook & Pettit, 2009). This assessment relied heavily on the assertions and representations of comScore and thus was explicitly described as "not an . . . audit" (Cook & Pettit, 2009, p. 2). Nonetheless, the report contains more methodological detail than can typically be found in publicly accessible documents; as well as a brief validation study in which the authors compare comScore data with other available data sources. Specifically, the authors compare comScore data on ad impressions with impression data obtained directly by the advertisers (finding variations of between 3–5% between the two data sources). In the end, the authors conclude with a very positive assessment of the comScore methodology and data (Cook & Pettit, 2009).

Of course, one of the greatest challenges in determining the accuracy and validity of any audience ratings system is finding any kind of meaningful baseline for comparison. When it comes to online ratings panels, there is one approach that is commonly relied upon. Specifically, as comScore's Gian M. Fulgoni (2003) has noted, "Unlike offline media, there is a reliable way to confirm the validity of sample-based online audience data: namely a comparison to site server logs, which measure machine-based visits" (p. 9). However, as Fulgoni (2003) also notes, this approach has, "historically, been a source of much heated debate between syndicated audience measurement suppliers and site operators, with syndicated suppliers generally showing lower audience estimates" (p. 9). Research has found discrepancies across the two approaches ranging from 35%– 350% (McDonald & Collins, 2007).

Reasons for the significant discrepancies include the fact that panel-based systems have tended to underrepresent work- and school-based Internet usage; the fact that web site visits can come from around the globe, though most panels are not global in nature; and the fact that server logs measure visits by individual computers/devices rather than individual people (Fulgoni, 2003). From this standpoint, the closer that panel-based estimates correlate with server-log estimates can be seen as reflective of the greater comprehensiveness of the panel being utilized (see Fulgoni, 2003, for an example of a comScore-conducted validation analysis comparing comScore's audience estimates for individual web sites with those of competing providers).

Of course, such approaches to assessing the validity of panel-based audience ratings services are somewhat limited in that, while they can facilitate comparative analyses of the size of the audiences measured by different analytical approaches, they cannot provide meaningful insights into the accuracy or validity of the demographic data obtained by panel-based ratings systems.

As should be clear, the limited analyses that are publicly accessible suggest persistent disagreements between the various methodological approaches to measuring audiences online. What is perhaps most striking about this state of affairs is the fairly recent observation by industry research professionals that "we know of no endeavors to attempt detailed reconciliations between the various methods, which might lead to an understanding and quantification of various areas of discrepancy" (Bennett et al., 2009, p. 165), and that, over time confusion, contradictions, and uncertainty, both within and across audience ratings systems have become worse, rather than better (Graves et al., 2010).

## 17.8 Statistical adjustment and modeling

Because most unweighted Internet audience panels are unrepresentative of the populations they purport to measure (as explained above), panel vendors use a variety of statistical adjustments and modeling techniques to try to correct these serious limitations. These techniques all amount to some form of statistical weighting and in some cases utilize a so-called "calibration panel" to assist weighting.

Calibration panels are ones that are formed by a probability sampling design (e.g., RDD or address-based sampling) and utilize more muscular and expensive recruitment protocols to try to yield a truly representative sample of the target population. These calibration panels are made up of members who agree to engage in the same Internet measurement systems in which the main measurement panel members engage. In theory, if the calibration panel is in fact representative of the target population, then the main panel can be modeled on the calibration panel, thereby making the main panel representative of the target population. However, we

are unaware of the existence of any publicly available detailed evidence that this approach is reliable and valid.

Another major problem in this area is that no detailed evidence has been reported to allow any outsiders to know with confidence whether any of the weighting that is used to adjust Internet measurement panels actually is successful in fixing them to be representative of their target population. However, it is unlikely that they are because of the nature of the factors that lead such panels to be unrepresentative in the first place. For all Internet measurement panels, there is a considerable amount of noncoverage or nonresponse that takes place when they first are built. Although the panel vendors can statistically adjust the panel to the demographic factors that are measured from the panel members and are available on the population, there is no reliable evidence of which we are aware that demonstrates that these demographic factors are in fact the characteristics that cause the lack of representation in these panels. Instead, it is likely that there are a host of unmeasured psychographic and lifestyle variables that account for why someone does or does not get sampled for an Internet measurement panel and for why someone who is sampled does or does not join the panel. Until panel vendors address these issues and release detailed evidence that (1) they have identified the key characteristics that cause their unweighted panels to be unrepresentative; and (2) they have used valid statistical means to take these key characteristics into account, it is our position that one cannot have confidence in the effect of any of the weighting schemes being deployed.

## 17.9 Representative research

The audience behavior and demographic data obtained via online ratings panels can, of course, be applied to a variety of both applied and academic research purposes. Advertisers and media buyers use the data to identify web sites that attract audiences that meet their desired audience size and demographic criteria. Web site operators use the data to assess the performance of their sites, and of individual pages within their sites. Such analyses can involve assessing performance patterns over time; and engaging in comparative performance analyses with competing web sites. Such analyses can be used to formulate and adjust strategy. Policy-makers will, on occasion, utilize such data to inform both economic and social policy concerns, such as assessing the competitive environment and determining which demographic groups do, and do not, have access to the Internet. Government agencies have utilized ratings data to assess usage patterns and audience demographics of government-operated web sites, as one of a variety of mechanisms employed to evaluate whether these sites are effectively meeting the population's information needs (Wood, 2005).

Online ratings panel data have also been utilized in academic research across a variety of fields, including communications, information science, and marketing. However, given the costs associated with obtaining access to such data, the use of such data for academic research purposes has been infrequent.

One important way that online ratings panel data have been utilized in academic research has been to develop and test theories of audience behavior. The bulk of the extant audience behavior research has been conducted within the context of traditional media such as television; and so one key question that has engaged researchers is whether patterns of audience behavior that have characterized traditional media are replicating themselves online.

Television audience researcher, James Webster, for instance, has been engaged in a line of research examining the extent to which patterns of audience duplication, fragmentation, and polarization that have come to characterize television viewing also characterize online

behaviors. Webster and Lin (2002), for instance, used Nielsen NetRatings data to examine whether patterns of audience distribution and audience duplication across web sites exhibited patterns found in the television context. In both cases, the authors found significant commonalities across television and online audience behaviors, leading them to conclude that "when users of the World Wide Web are seen as a mass audience, they exhibit straightforward, law-like behaviors similar to those of older media" (Webster & Lin, 2002, p. 8).

This line of inquiry is extended in Webster and Ksiazek's (2012) analysis of audience duplication *across* television channels and web sites; this time utilizing a Nielsen "convergence panel" of individuals who simultaneously had their web browsing and television viewing monitored. Their results indicate levels of audience duplication across television channels and web sites that would seem to call into question commonly-raised concerns about media fragmentation leading to audiences segmenting themselves into narrowly defined, homogeneous enclaves (Webster & Ksiazek, 2012).

Similar concerns underlie Tewksbury's (2005) analysis of online news audiences. Utilizing Nielsen NetRatings data, Tewksbury (2005) addresses the political dimension of online media consumption, in terms of whether individual news sites are extremely specialized in terms of the types of audiences (in terms of demographics) they attract, and in terms of the type of content that is produced and consumed. The results indicate high levels of specialization across both audience demographics and content production and consumption (Tewksbury, 2005).

Chiang, Huang, and Huang (2010) look specifically at the well-known "double jeopardy" effect (a pattern in which there is a strong correlation between audience size and audience loyalty), in an effort to determine whether this pattern is prevalent in patterns of online audience behavior. Utilizing ratings panel data provided by the Taiwanese audience measurement firm InsightXplorer Limited, the authors find strong evidence that the double jeopardy effect is as prominent in online audience behavior as it has been in the behavior of traditional media audiences (Chiang et al., 2010).

Marketing researchers Danaher, Mullarkey, and Essegaier (2006) utilized online panel data to explain patterns in the *duration* of audiences' visits to individual web sites. Given the widely acknowledged importance of the duration of audience visits to maximizing audiences' exposure – and potential response – to online advertisements, the authors sought to identify those factors that affect visit duration. Utilizing Nielsen NetRatings data, the authors found a wide range of significant predictors of visit duration, including: demographic characteristics such as age (older audiences = longer visit duration) and gender (female = longer visit duration); as well as content characteristics (auction, entertainment, and news sites = longer visit duration). Importantly, the authors also found that the amount of advertising content is significantly negatively related to visit duration (Danaher et al., 2006).

## 17.10 The future of Internet audience measurement

It is important to emphasize that today, almost 20 years after the birth of Internet audience ratings systems, most of the stakeholders involved remain highly dissatisfied with the nature of the information available to them. As one recent trade press account began, "Online measurement is broken. That, few would dispute" (Neff, 2011, p. 8). This same account details recent efforts by major advertisers, trade organizations, and media companies to establish acceptable standards for the measurement of online audiences and the metrics to be reported. Called "Making Measurement Make Sense," this initiative in many ways reflects how little

progress has been made over the past 20 years in terms of measurement firms being able to satisfy the demands of their client base. It is also a reflection of the enormous complexity and challenges associated with the task of producing online audience ratings. It seems reasonable to question whether a system that genuinely satisfies all of the relevant stakeholder groups is even achievable under the challenging conditions of the online media environment.

The key trend going forward appears to be one in which online panels no longer function as a stand-alone source of ratings data. The current state of the art in Internet ratings involves a hybrid approach that integrates panel-centric and site-centric data. In 2009, for instance, comScore (2009) launched Media Metric 360, a service that derives web site ratings from what it terms a "panel-centric hybrid" that combines the ability of site-centric data to capture audience exposure in the long tail with the demographic data that can best be obtained via traditional panels (for more detail, see Pellegrini, 2009). Such an approach provides a means of extending the reach of an Internet ratings system further into the tail than can be achieved via a pure reliance on panels. However, such an approach does not completely "solve" the problem of the long tail (discussed above).

An additional methodological approach, therefore, is likely to become more fully integrated into the production of Internet audience ratings as these measurement services continue to evolve. Specifically, "direct" measurement of Internet traffic data obtained from ISPs is also increasingly being integrated into the production of online audience ratings (by firms such as Quantcast). Thus, according to some industry professionals, for "the immediate future and beyond, we'll be living in a three-sourced world for Internet audience estimates: Panel, server and direct" (Bennett et al., 2009, p. 166).

Finally, probably the biggest challenge in Internet audience estimates is the proliferation of the number of devices per user. It is very common that in developed countries the same people access the web from their desktop computer at work, a laptop computer, a tablet, and a smartphone. Measuring the "whole picture" is extremely complex and technically challenging, not to mention privacy requirements. Advertisers are also interested in learning about the interaction of traditional media such as TV and print media with Internet advertisements. Measuring such exposure within the same households, for example, adds another layer of complexity to an already fragmented Internet browsing device behavior. This is, however, the direction that Internet rating companies are pursuing with different measurement solutions.

## References

Abraham, M. M. (2012, April 23). comScore plants 3 millionth tree as part of "trees for knowledge" program. Retrieved July 1, 2013, from: http://www.comscore.com/Insights/Blog/comScore_Plants_3_Millionth_Tree_as_Part_of_Trees_for_Knowledge_Program.

Anderson, C. (2008). *The long tail: How the future of business is selling more of less* (Updated ed). New York: Hyperion.

Arbitron. (2012a). Media tracker. Retrieved August 24, 2012, from: http://www.zokem.com/products/media-tracker/.

Arbitron. (2012b). Arbitron mobile infographics. Retrieved August 24, 2012, from: http://www.zokem.com/arbitron-mobile-infographics/.

Arini, N. (2012a, March 19). Measuring modern media consumption. Retrieved August 24, 2012, from: http://adwordsagency.blogspot.co.uk/2012/03/measuring-modern-media-consumption.html.

Arini, N. (2012b) http://www.google.com/landing/screenwisepanel/.

Bennett, D., Douglas, S., Rogers, B., & Broussard, G. (2009). Digital research – online measurement too many numbers – too many different relationships. Paper presented at

the Worldwide Readership Research Symposium, Valencia, Spain. Retrieved July 2, 2012, from: http://www.printanddigitalresearchforum.com/online-measurement-%E2%80%93-too-many-numbersrelationships.

Bermejo, F. (2007). *The Internet audience: Constitution & measurement*. New York: Peter Lang.

Bermejo, F. (2009). Audience manufacture in historical perspective: From broadcasting to Google. *New Media & Society*, *11*, 133–154.

Blumberg S. J., & Luke J. V. (2013) *Wireless substitution: Early release of estimates from the National Health Interview Survey, July–December 2012*. National Center for Health Statistics. June 2013. Retrieved June 29, 2013 from:http://www.cdc.gov/nchs/data/nhis/earlyrelease/wireless201306.pdf.

BusinessWire (2011, September 13). Nielsen releases first mobile media rankings based on audience measurement data from Android smartphone meters. Retrieved August 22, 2012, from: http://www.businesswire.com/news/home/20110913005752/en/Nielsen-Releases-Mobile-Media-Rankings-Based-Audience.

Chiang, I., Huang, C., & Huang, C. (2010). Traffic metrics and web 2.0-ness. *Online Information Review*, *34*, 115–126.

Coffey, S. (2001). Internet audience measurement: A practitioner's view. *Journal of Interactive Advertising*, *1*, 10–17.

Compete (2011). Overview of data methodology and practices. Retrieved August 19, 2012, from: http://media.compete.com/site_media/upl/img/Compete_Data_Methodology_3.pdf.

comScore (2009). comScore announces Media Metric 360: The next generation of global digital audience measurement. Retrieved August 28, 2012, from: http://www.comscore.com/Press_Events/Press_Releases/2009/5/comScore_Announced_Media_Metrix_360.

comScore (2012). Investor FAQs. Retrieved August 28, 2012, from: http://ir.comscore.com/faq.cfm.

comScore (2013). comScore brings big data to mobile and multi-platform analytics. Retrieved February 22, 2013, from: http://www.comscore.com/Insights/Press_Releases/2013/2/comScore_Brings_Big_Data_to_Mobile_and_Multi-Platform_Media_Analytics.

comScore (n.d.). GSMA mobile media metrics. Retrieved February 22, 2013, from: http://www.comscore.com/Products/Audience_Analytics/GSMA_Mobile_Media_Metrics_MMM.

Connected Vision. (2011, July 8). Google and Kantar develop measurement panel. Retrieved August 24, 2012, from: http://informitv.com/news/2011/07/08/googleandkantar/.

Cook, W. A., & Pettit, R.C. (2009, April 6). comScore Media Metrix U.S. Methodology. An ARF research review. Retrieved July 1, 2013, from: http://thearf-org-aux-assets.s3.amazonaws.com/downloads/research/comScore-RReview.pdf.

Danaher, P. J., Mullarkey, G. W., & Essegaier, S. (2006). Factors affecting web site visit duration: A cross-domain analysis. *Journal of Marketing Research*, *43*, 182–194.

Earle, B., Shaw, K., & Moy, C. (2011). The GfK NOP Media Efficiency Panel. Retrieved August 24, 2012, from: http://www.mrs.org.uk/article/item/356.

FAST/ARF (2000). Principles for online media audience measurement. Retrieved August 2, 2012, from: http://s3.amazonaws.com/thearf-org-aux-assets/downloads/research/ARF-FAST-Principles-2000.pdf.

Fitzgerald, K. (2004, March 15). Debate grows over Net data. *Advertising Age*, pp. 4, 78.

Flynn, L. (1995, May 29). In search of Nielsens for the Internet. *New York Times*. Retrieved July 25, 2012, from: http://www.nytimes.com/1995/05/29/business/in-search-of-nielsens-for-the-internet.html?pagewanted=all&src=pm.

Fulgoni, G. M. (2003, June). Measuring consumers' Internet behavior. Paper presented at the ESOMAR/ARF Worldwide Audience Measurement Conference, Los Angeles, CA.

Gluck, M. (2011). State of mobile measurement. Paper prepared for the Interactive Advertising Bureau. Retrieved August 2, 2012, from: http://www.iab.net/media/file/Mobile-measurement-paper-final-draft2.pdf.

Goosey, R. W. (2003, June). Defining the gold standard for user centric global online audience measurement. Paper presented at the ESOMAR/ARF Worldwide Audience Measurement Conference, Los Angeles.

Haering, H. (2002, June). Monitoring Internet audience measurement panels. Paper presented at the ESOMAR/ARF Worldwide Audience Measurement Conference, Cannes, France.

Holden, R., Heng, L. M., Lavrakas, P. J., Bell, S., & Flizik, A. (2006). Personal contact and performance-based incentives to raise long-term panel compliance and to reduce missing data. Paper presented at the 61st annual conference of the American Association for Public Opinion Research Conference, Montreal.

Iannacchione, V. G. (2011). The changing role of address-based sampling in survey research. *Public Opinion Quarterly*, *75*, 556–575.

Interactive Advertising Bureau (2009). Audience reach measurement guidelines version 1.0. Retrieved July 24, 2012, from: http://www.iab.net/media/file/audience_reach_022009.pdf.

Kok, F., Appel, M., Verhulst, E., & Pellegrini, P. (2011). Internet audience measurement now – sites, content and campaigns – anywhere, anyhow. Paper presented at the Print and Digital Research Forum, San Francisco. Retrieved August 19, 2012, from: http://www.printanddigitalresearchforum.com/internet-audience-measurement-now-sites-content-and-campaigns-anywhere-anyhow.

Lavrakas, P. J., Dennis, M. J., Peugh, J., Shand-Lubbers, J., Lee, E., Charlebois, O., & Murakami, M. (2012). An experimental investigation of the effects of noncontingent and contingent incentives in recruiting a long-term panel: Testing a Leverage Salience Theory hypothesis. Paper presented at the 37th Conference of the Midwest Association for Public Opinion Research. Chicago, IL. Retrieved from: http://www.mapor.org/confdocs/absandpaps/2012/2012_papers/5C4_Lavrakas_paper.pdf.

Levine, D., & Finkle, J. (2011, August 23). Exclusive: Privacy lawsuit targets comScore. *Reuters*. Retrieved February 22, 2013, from: http://www.reuters.com/article/2011/08/23/us-comscore-lawsuit-idUSTRE77M76O20110823.

McDonald, S. & Collins, J. (2007). Internet site metrics and print media measurement – from dissonance to consonance? Paper presented at the Worldwide Readership Symposium, Vienna. Retrieved August 19, 2012, from: http://www.gfkmri.com/PDF/WP%20Internet%20Site%20Metrics%20and%20Print%20Media%20Measurement%20From%20Dissonance%20to%20Consonance.pdf.

Media Rating Council. (2007). A guide to understanding Internet measurement alternatives. Retrieved July 1, 2013, from: http://mediaratingcouncil.org/MRC%20POV%20General%20Internet%20080307.pdf.

Napoli, P. M. (2003). *Audience economics: Media institutions and the audience marketplace*. New York: Columbia University Press.

Napoli, P. M. (2011a). *Audience evolution: New technologies and the transformation of media audiences*. New York: Columbia University Press.

Napoli, P. M. (2011b). Ratings and audience measurement. In V. Nightingale (Ed.), *The handbook of media audiences* (pp. 286–301). Chichester: Wiley-Blackwell.

Napoli, P. M., & Obar, J. (2013). Mobile leapfrogging and digital divide policy: Assessing the limitations of mobile Internet access. New America Foundation White Paper. Retrieved July 1, 2013, from: http://newamerica.net/sites/newamerica.net/files/policydocs/MobileLeapfrogging_Final.pdf.

Neff, J. (2011, Sept. 19). Coca-Cola, AT&T and others out to reinvent web measurement. *Advertising Age*, p. 8.

Nielsen. (2009, September 22). Nielsen in a relationship with Facebook. Retrieved from: http://www.nielsen.com/us/en/newswire/2009/nielsen-in-a-relationship-with-facebook.html.

Nielsen. (2011, May 5). Connected devices: how we use tablets in the US. Retrieved from: http://blog.nielsen.com/nielsenwire/online_mobile/connected-devices-how-we-use-tablets-in-the-u-s/.

Nielsen (2012a). The industry benchmark for Internet audience measurement. Retrieved July 5, 2012, from: http://nielsen.com/content/dam/nielsen/en_us/documents/pdf/Fact%20Sheets/NetView_US.pdf.

Nielsen (2012b). Press center: Press FAQs. Retrieved August 22, 2012, from: http://www.nielsen-online.com/press.jsp?section=pr_faq.

Nielsen (2012c). Smartphone study: FAQs & contacts. Retrieved August 22, 2012, from: https://mobilepanel.nielsen.com/enroll/help.d0?l=en_us&pid=1.

Nielsen. (2012d). *The cross platform report*: *Q2 -2012 - US*. Retrieved July 1, 2013, from http://www.nielsen.com/content/dam/corporate/us/en/reports-downloads/2012-Reports/Nielsen-Cross-Platform-Report-Q2-2012-final.pdf.

Ofcom (2013). Ofcom technology tracker Wave 3, 2012. Retrieved from: http://stakeholders.ofcom.org.uk/binaries/research/statistics/2013jan/wave3.pdf.

Opidee, I. (2012, May 8). comScore launches smartphone metrics services. *FolioMag*. Retrieved August 22, 2012, from: http://www.foliomag.com/2012/comscore-launches-smartphone-metrics-service.

Pellegrini, P. (2009). Towards panel-centric hybrid measurement: Enhancing online audience measurement. Paper presented at the Worldwide Readership Research Symposium, Valencia, Spain. Retrieved August 12, 2012, from: http://www.printanddigitalresearchforum.com/papers/921.pdf.

Pew Internet. (2012). Adult gadget ownership over time (2006–2012). Retrieved August 26, 2012, from: http://pewinternet.org/Trend-Data-%28Adults%29/Device-Ownership.aspx.

Raine, L. (2012, September 11). Smartphone ownership update: September 2012. Retrieved from: http://pewinternet.org/Reports/2012/Smartphone-Update-Sept-2012.aspx.

Raine, L., & Duggan, M. (2012, December 27). Ebook readings jumps, print book reading declines. Retrieved July 1, 2013, from: http://libraries.pewinternet.org/2012/12/27/e-book-reading-jumps-print-book-reading-declines/.

Reynolds, J. (2012, September 24). Facebook partners with Datalogix for ad tracking project. Retrieved from: http://mediaweek.co.uk/news/1151339/Facebook-partners-Datalogix-ad-tracking-project/.

Smith, A., & Boyles, J. L. (2012, July 17). The rise of the "connected viewer". Retrieved July 1, 2013, from: http://www.pewinternet.org/~/media//Files/Reports/2012/PIP_Connected_Viewers.pdf.

Spangler, T. (2012, February 13). NBCU enlists Google, comScore to track multiscreen Olympics viewing. Retrieved August 24, 2012, from: http://www.multichannel.com/article/480514-NBCU_Enlists_Google_ComScore_To_Track_Multiscreen_Olympics_Viewing.php.

Steinberg, B. (2012, February 13). NBC Universal's new Olympics challenge: Screen-jumping. Retrieved August 24, 2012, from: http://adage.com/article/media/nbc-universal-s-olympics-challenge-screen-jumping/232676/?page=1.

Story, L. (2007, October 22). How many site hits? Depends who's counting. *New York Times*. Retrieved July 22, 2012, from: http://www.nytimes.com/2007/10/22/technology/22click.html?_r=1&pagewanted=all.

Tewksbury, D. (2005). The seeds of audience fragmentation: Specialization in the use of online news sites. *Journal of Broadcasting & Electronic Media*, *49*, 332–348.

Webster, J. G., & Ksiazek, T. B. (2012). The dynamics of audience fragmentation: Public attention in an age of digital media. *Journal of Communication*, *62*, 39–56.

Webster, J. G., & Lin, S. (2002). The Internet audience: Web use as mass behavior. *Journal of Broadcasting & Electronic Media*, *46*, 1–12.

Webster, J. G., Phalen, P. F., & Lichty, L. W. (2005). *Ratings analysis: The theory and practice of audience research*. Mahwah, NJ: Erlbaum.

Wood, F.B. (2005). Use of Internet audience measurement data to gauge market share for online health information services. *Journal of Medical Internet Research*, *7*. Retrieved August 28, 2012, from: http://www.jmir.org/2005/3/e31/.