Chapter 17
Digital Metrics: Getting to the Other 50 Percent

Michelle R. Nelson
University of Illinois at Urbana-Champaign, USA

Helen Katz
Starcom Mediavest Group, USA

ABSTRACT
This chapter reviews measures of advertising effectiveness in research and practice from the pre-digital to the digital era. A focus on efficacy and ethics in terms of measurement and consumer privacy issues associated with collecting, monitoring and learning from digital metrics is discussed. Research questions related to persuasion knowledge and digital privacy are posed.

‘Half the money I spend on advertising is wasted; the trouble is I don’t know which half.’ John Wanamaker, (attributed) U.S. department store merchant (1838–1922)

INTRODUCTION
These famous (or infamous) words are still cited today as marketers try to determine how to reach their desired target audiences with the right message in the right medium at the right time to inspire those consumers to purchase (more of) the marketers’ brands. Media measurement has undergone vast transformation since Wanamaker’s time. For example, in 1906 radio broadcasters simply asked, “Is anybody out there?” suggesting listeners should write a letter to inform them. More than one hundred years later, planners have access to cross-measurement and fusion media data collecting media habits and linking them to actual buyer behavior (Sass 2007). Yet, one of the most basic research questions—who is the audience—is still being discussed (Webster, Phalen & Lichty 2006).

Today, marketers want and need to know much more than that, however. With the explosion in media choices, and the desire for consumers to have greater control over their media and ad exposure, marketers are challenged to gain more precise and accurate information, not only on who is the audience, but where they are and how receptive
they might be to ad messages. *Digital metrics* (audience measurement of digital media) present a great opportunity to collect, gather and use a remarkable amount of audience information. Such information offers advertisers new micro-targeted means of reaching audience members. Yet, with great data also comes great responsibility for the ethical management of that information to protect consumer privacy. In this chapter, we review the transition from old to new metrics, discuss what works today, and what is needed tomorrow in terms of measurement and consumer privacy issues associated with collecting, monitoring and learning from digital metrics.

The IDC Digital Marketplace Model and Forecast predicts that total worldwide Internet advertising will be as much as $51.1 billion in 2012. Although still a relatively small proportion of overall ad spending, the number is growing faster than all other forms of media. Despite this industry growth, as well as an increasing focus among lawmakers and regulators on behavioral targeting and online advertising, research has yet to fully address these important issues. We provide an overview of the key concepts along with a number of future research questions in the hopes of spurring more attention in this area.

**METRICS: ADVERTISING EFFECTIVENESS THEN: THE EYEBALLS (AND EARS) HAVE IT**

Measurement of advertising effectiveness—whether digital or not—should conform to advertising objectives (Li & Leckenby 2007). So if the objective of an advertising campaign is to increase brand awareness, then a direct response behavioral metric, such as how many times someone clicked on a banner ad, may not be the most appropriate measure. Indeed, our contemporary thinking about advertising objectives and measurement goes back more than one hundred years. The earliest model of advertising effectiveness, created in 1898 by Elmo St. Lewis, focused on ‘attention, interest, desire, and action’ (AIDA) (Barry 1987). Subsequent academic models of effectiveness, such as the “hierarchy of effects” model by Lavidge and Steiner (1961), also focused on cognition (thinking), affection (liking) and conation (behavior). Although Ray (1973) found some evidence for the existence of cognition, affection, and conation, there have been critiques of such hierarchical models of advertising effectiveness (see Robertson 1970; Weilbacher 2001). Measures of advertising effectiveness in academic studies have largely followed these hierarchies—including advertising recall or recognition (for cognition), attitude toward the ad and ad liking (for affection) and conation (usually purchase intent).

Real-world metrics of traditional media have also relied on advertising attention and recall, employing active techniques of measurement by asking people questions. For example, in 1923 Dr. Daniel Starch began analyzing print advertisements by interviewing people and asking them whether the ad was “noted” (reader remembers seeing ad), “associated” (reader remembers seeing name of advertiser) and “read most” (reader actually read at least half of the ad). Similarly, for broadcast media, the first ratings service (“Crossley ratings”) telephone interviewed potential radio listeners in the mid-late 1930s asking them to recall their own radio listening during the past 3-6 hours. However, even those early researchers noticed that people did not or could not always remember what they had listened to. This active measurement technique was flawed. In addition, the metric focused on media exposure and not advertising exposure. A competitor at that time, pollster George Gallup, devised a method that concentrated not on recall but on current listenership (“coined telephone coincidental”). Questions focused on what programs audiences were listening to at the time of the call, what station they were tuned to, and the name of the sponsor of the program. Demographic information (age, gender) was also collected. This method was instituted into a syndicated ratings practice.
by Clark-Hooper in 1934. The service began reporting audience shares, percent of listeners and demographic information—thus “by the end of the 1930s the basic pattern of commercial audience research for broadcasting was set” (Webster et al. 2006, p. 98). Subsequent measures included “aided recall” where respondents were given a roster of programs and asked in a personal interview which programs they listened to and in radio diaries to write down radio listening.

All of these measurement techniques required audience members to remember what media they were exposed to either in the past or the present. Such methods have been criticized because people either do not remember or they over-report or under-report certain kinds of programs (Webster et al. 2006). However, more passive forms of measurement were developed to circumvent some of these problems. These are reviewed next.

**METRICS: TECHNOLOGY DRIVES TRANSITIONS**

So why the changes? In a word - Technology. The shift to bits and bytes from dots and spots has resulted in many changes not only in how media audiences are being measured and reported, but also in marketers’ expectations of what these media can do for them.

It has been a gradual but ongoing transition. The first shifts were seen as far back as 1942 when the A.C. Nielsen Company, which had acquired and then perfected an audiometer device, used it to record when the radio set was turned on (Webster et al. 2006). The company launched the Nielsen Radio index and combined it with an inventory of each household’s pantry (purchase data). Then in the 1960s, Nielsen moved its television audience measurement away from paper diaries to TV set meters (a technology that still relies on). That made the measurement electronic, removing some of the ‘burden’ on consumers to remember and relay their media (and advertising) exposure.

It was not until the 1980s, and under the threat of competition, that Nielsen introduced additional electronic measurement in the form of the People Meter. This device, similar to a TV remote control, enabled viewers in the national Nielsen sample to punch in and out when they were watching TV, at the time they were doing so, rather than recording on paper every 15 minutes of their TV day. Today, there are 18,000 homes participating in the national People Meter service, with several hundred People Meters also measuring viewing activity in the top 25 local markets.

In the 1990s, Arbitron went down a similar, and in some ways more advanced, technological path in the quest to enhance audience measurement metrics. Its introduction of the Portable People Meter (PPM) removed the need for any kind of button pushing or active participation. By wearing a pager-like device that automatically picks up radio (and TV) signals embedded in the media content, a panelist can provide that data passively to Arbitron. But it is the introduction, growth, and development of the Internet as an advertising medium that has truly changed the metrics involved in audience measurement.

**THE INTERNET: ADVERTISING AND METRICS**

By virtue of the fact that the Internet is able to passively collect every click of the computer mouse, to follow the user wherever they go on the Web, and to see explicitly which ads are clicked on, the world of audience measurement has been decisively changed. In traditional media forms, what has usually been measured is the audience’s ‘opportunity to see’ an ad—their exposure to a TV program or radio quarter hour or magazine or newspaper issue. With the Internet (and digital measurement in general), the greater level of data granularity enables marketers to capture exposure to the ad within the medium. What they are measuring and how they measure depends to some
extent on the advertising format. Each of these will be reviewed next within a general discussion of Internet ad measurement.

Advertising formats have changed since the first simple banner ads were created in 1997, which provided a billboard, or display ad similar to a static newspaper image. As video, sound, and movement were added, the ads became far more dynamic, and were called rich media. But perhaps the biggest change has been the growth of search advertising, using keywords. In 2008, more than 33% of ad spending worldwide was allocated to keyword ads, 20% to display ads, and 19% to classified ads (“Internet ad growth,” 2008).

**Keywords**

Under the name “search marketing,” keyword ads are also referred to as Textual Ads (short textual messages usually marked as sponsored links) (Chakrabarti, Agarwal, & Josifovski 2008). They work in two ways: (1) sponsored search/paid search where ads are placed on pages on a web search engine based on the search query. In this case, all the major search engines (e.g., Google) act as search engine and advertising agency combined; (2) contextual advertising or context match where ads are placed within a generic, third-party web page (usually through an ‘ad network’ intermediary). As the name suggests, these ads are placed according to the semantic similarity of the content and the ads (key word matches) as well as click-through feedback.

The two main ways to measure Internet ad effectiveness are by click through rates and conversion rates. **Click–through rates (CTRs)** are obtained by dividing the number of users who clicked on an ad on a web page by the number of times the ad was delivered (impressions), while **conversion rates** are the percentage of search-generated visitors who make a purchase or answer a call to action. Although these numbers are typically low in sponsored search (about 1-2%, the same response rate, on average, as direct mail), other potential benefits such as cross-selling and advertising brand-specific keywords (Ghose & Yang 2008) are offered. For example, retailers can pair the searched-for product with other products that sell well on the same web site and ‘direct’ searchers to other popular products.

**Classified Ads**

Classified ads–where individuals find the right buyers for their products–have found new forms on the Internet beyond local newspapers (Diaz 2007). For the first time in almost 100 years, newspapers are losing revenue from classified advertising to specialized classified sites such as Craigslist or classifieds on Facebook. Indeed, online traffic to such sites has grown to 42.2 million unique visitors a month, according to March 2009 figures (“Online classified,” 2009). Forty-five percent of all Internet users have used online classified sites. As a result, the classified industry is in a time of transition–but buyers and sellers themselves have more options to find one another.

**Display Ads**

Display ads, (often called banner ads) come in many forms and sizes in digital media. They can be text, graphics, static or animated, and even interactive, with more than 14 different sizes (Li & Leckenby 2007). Rich media display ads allow users to interact with the content or even with marketers or other customers. For example, Volvo piped Twitter into its ad units as a way to respond directly with customers (Morrisey 2009). Advertisers are also experimenting with live video and 3D technology within its display ads. Visa’s global campaign featured real-time scenes from cities around the world in their display ads (Morrisey 2009).

Although banner ads were one of the first forms of advertising on the Internet, their share of Internet advertising budgets has fallen in recent years from 1999 when they accounted for 56%
of the market (IAB) to just 20% in 2008, even as the actual money spent on display ads has risen. In 2008, growth in elaborate display ads slowed down (Clifford & Helft 2008). This was due, in part, to the economic downturn, but also because ad dollars have shifted into search-based ads, reflecting that “advertisers are becoming more performance, ROI-focused,” according to an analyst at Jeffries & Company.

Display ads can be measured in a number of ways. Traditionally, they were viewed as a direct response tactic, thus click-through rates were the primary measure. However, perhaps due to falling click-through rates (i.e., from their high of a 3% in the 1990s to less than 0.5% in recent years (Li & Leckenby 2007), additional measures are now employed, such as impressions, which can show display ad views—i.e., the number of times an ad is viewed, where they are viewed, demographics of those exposed to ads, and frequency. A “cost-per-engagement” pricing model only charges marketers when the consumer interacts with the ad (Steel 2009). The use of impressions as a digital metric also enables media planners to make more direct comparisons to traditional media forms, such as television or print, where the same metric can be calculated.

Within academic research, effectiveness of display ads has been addressed in both experimental (e.g., Li and Bukovac 1999) and field-based (e.g., Rosenkrans 2009) studies. This relatively small body of research has demonstrated that increased interactivity (e.g., Chandon, Chtourou, and Fortin 2003), banner size (e.g., Li and Bukovac 1999), color-contrast (Dreze and Husherr 2003), and rich media capability (Rosenkrans 2009) contribute to higher click-through rates. However, click-through is a behavioral metric that requires conscious attention on the part of the audience. With increasing advertising clutter, it is likely that audiences do not or cannot attend consciously (or click through) all of the advertising on a web page. Yet, the display ad may still impact the audience in a manner not captured by click-through rates. Indeed, an early study by Briggs and Hollis (1997) found that banner ads were able to increase ad awareness, brand perceptions, and favorable attitudes—even without click-throughs, thus pointing out the potential for image building or brand equity through display ads. Further, in a more recent experimental study, Yoo (2008) manipulated various levels of consumer attention and then measured recall, attitudes toward the advertised brands, and placement of advertised brand into consideration sets. The findings suggest that even when consumers did not consciously process the ad and could not explicitly recall the brand, the ads had some positive effects on consumers’ brand attitudes. Further, those with exposure to the brands but not conscious exposure were more likely to place the advertised brand into their consideration sets than those who were not exposed to the brand at all. Yet despite some promise of branding offered in a handful of academic studies, the majority of marketers still question the viability of brand-building online (Morrissey 2009). Such beliefs relate to the metrics most commonly used in industry to measure effectiveness. Results of a 2009 survey of top marketers by Forbes.com found that only 31 percent of marketers regarded brand building as a viable metric for digital advertising as compared with 51% who used click-through rates (Morrissey 2009).

Some studies have also compared web-based measures with more traditional measures (e.g., attitude toward the brand). The results are, however, not conclusive. For example, a study by media agency Starcom, behavioral targeting network Tacoda, and ComScore—revealed that there were no correlations/statistical relationships between display ad clicks and brand metrics (attitude toward brand/click-thru rate) (ComScore, 2008). Further, the study showed that “heavy clickers” (those who accounted for 50% of all display clicks) made up only 6% of those online; thus, clickers were by no means representative of the general or online public. However, another study conducted by ComScore for the pharmaceutical industry
showed that the impact of banner (display) ads, search marketing and visits to the branded Web site all resulted in increased brand awareness and favorability. Such positive effects were especially true for prospective users of the products (‘prospects’). Results showed greater aided brand awareness when audience members interacted with a rich media ad (The Center for Media Research, 2008). Clearly, given the inconclusive evidence, more research is needed in this area. Ideally, the research could employ theoretical models of audience involvement and new understandings of implicit persuasion from psychology (e.g., Petty et al. 2009) to better understand measures such as engagement and impressions and discern how the various metrics fit together.

MORE METRICS: DATA COLLECTION & TARGETING

Beyond click-throughs and impressions, other metrics that are routinely captured for Internet audiences include:

- **Time spent viewing**: how much time was spent on a particular website or web page
- **Reach and frequency**: traditional media metrics showing the percent of a target group reached by a website and/or campaign. The reach represents the unduplicated audience for that site/page, while frequency shows how often that group is reached by the site/page
- **Behavioral target activity**: viewing behavior/activity among a group defined based on their web activity (as opposed to demographics)
- **Retention rate**: percent of a group visiting a given website last month who also visited the site this month.
- **Conversations**: amount and type of ‘talk value’ about an advertisement
- **Registrations**: number of people who sign up or register on the website as a result of an ad
- **Conversion Rates**: percent of those who clicked on an ad that results in a sale

All of these metrics vary depending on how they are defined. For most traditional media, they either rely on active reporting of behavior (such as the Nielsen people meter, where panelists have to press a button every time they start and stop watching TV), or on consumer recall (“do you remember seeing this particular magazine in the past 30 days?”). For the Internet, the measurement is passive. For syndicated measurement companies, once someone has agreed to be part of the panel, all the company needs to do is load special software onto panelists’ computers and all computer activity is passively captured. Websites can collect most of that even without panelist permission. This is called Behavioral Tracking or Behavioral Targeting and it provides advertising to Internet users based on their Web surfing habits.

The way that the behavior is captured is through the use of **cookies**. Originally referred to as “magic cookies”–named after tokens with mystical powers in role-playing games (Wildstrom 1996) - they are “placement of small text files on a consumer’s hard drive that are then offered back to the Web site during subsequent visits by the consumer” (Miyazaki 2008, p.20). The Web site itself provides the cookie, and then the browser installs the cookie on the computer’s hard drive (Davidson 2007). A range of data can be collected from behavioral data tracking user movements—including time, duration, and sequence of movements; user information such as demographics; passwords; and media tracking—(i.e., how many times a banner ad appeared). The total collection of data that is captured on the computer is referred to as **clickstream data** (Erickson 1996). The Internet measurement companies tie that data to audience behavior—who has been exposed, and
for how long. That, in turn, can be classified into consumer behaviors—any shopping or commerce conducted over the web. And they can deep-dive into online-specific behaviors, including instant messaging, email, gaming, or streaming media (audio and video). The data are then projected from the panel to the total online universe.

Unlike spyware, which offers no real value to consumers (Davidson 2007), cookies can be useful. For example, because certain types of information are stored on Web sites, consumers do not need to re-enter personal data (such as a mailing address) when visiting certain Web sites, their password is stored in a cookie - such as the New York Times web site (Berg 1997). Or they are able to pay for items in a virtual shopping cart at the end of their online shopping (rather than separately for each item). Interestingly, even the legal definition of cookies focuses on positive aspects, “Cookies are computer programs commonly used by Web sites to store useful information such as user names, passwords and preferences, making it easier for users to access Web pages in an efficient manner” (see Davidson 2007, p. 447). However, the downside occurs when the data collected are more than a consumer wants to share or when the information is being used for a purpose for which s/he did not consent.

From a marketer’s standpoint, the behavioral information gained from cookies can be used to create user profiles for more effective segmenting and targeting in advertising and products, and the data do not usually provide any personally identifiable information. Segmenting, the “process of dividing the market into more manageable submarkets or segments” (Urban 2004), is typically based on consumer factors, product factors or media factors (see Rodgers, Cannon & Moore 2007). Once consumers have been segmented into demographic or passion groups (e.g., dog lovers)—advertisers can use the information to make sure advertising gets the ‘right’ message to the ‘right consumer’—reducing the financial waste of paying to reach everyone (Beane & Ennis 1987).

This is where targeting comes in. Targeting, “the process of delivering content or ads to segments or visitors based on their known attributes” (Phillips 2008), allows advertisers to tailor message content to fit particular aspects of the consumer when and where they are the most receptive. Tailored messages are attended to (Pechmann & Stewart, 1990), more persuasive (Ariely 2000), better liked (Kern et al. 2008) and show a greater chance of being ‘clicked on’ (Chatterjee, Hoffman & Novak 2003) than general messages.

What digital metrics can provide, beyond the traditional ways to segment and target audiences (e.g. demographics - age, gender, ethnicity), is the ability to do behavioral targeting - the ‘technology and process in which an ad or content is shown to a visitor based on their past actions and behaviors’ (Phillips 2008). For example, advertising practitioners are already creating multiple versions of messages for different audience members (up to 20,000 versions of an ad for a single brand; see Story 2007, also Morrisey 2009). The goal is a more personalized message for each customer.

Personalization is a “specialized flow of communication that sends different recipients distinct messages tailored to their individual preferences or characteristics” (White et al. 2008, p.40). Personalization in the broader online world can be based either on explicit data collection (demographics, product ratings, opt-in information provided by consumer) or implicit data collection (inferred about user—based on cookies that collect information related to search queries, purchase history/ browsing history; Cranor 2003). For example, if a consumer is looking for airline tickets to San Francisco online, an ad may appear for spas in the area or San Francisco sights or restaurants. Personalization, however, can also be profile-based— whereby cookies are used to recognize returning visitors. When a consumer visits Amazon.com and the website welcomes him by name and even offers recommendations based on his past purchase behavior, it does so by analyzing his prior Internet activity as well as the personal
information it has stored from his previous visits to that site. Such personalization offers relevant information to benefit the consumer; however, the line between “personal enough to be useful” and “so personal it freaks me out” is a fine one (Mediapost, July 28, 2008). For example, Facebook’s personalization included a travel ad directed to a media blogger that read, “Hey Jew…”, which was promoting “adventure travel that’s worth the schlep!” (Berkowitz 2008) and another ad was targeted to an engaged woman that read, “Do You Want to be a Fat Bride?” (Beckman 2008). While such ads can certainly capture attention, it is questionable as to whether they are effective or offensive.

Indeed, this level of personalization can have a detrimental effect when the message is deemed “too personal.” An experimental study by White et al. (2008) showed that consumers had a negative response to ‘highly personalized’ e-mail solicitations, especially when the offer in the message was not directly related to the personal information offered (i.e., personalization was not justified). For marketers, then, the challenge is to use information to target (by media) and by message—with just the right combination of relevance and personalization.

The availability of all these data has transformed the way that marketers think about media. For once metrics became available that could link consumers’ exposure to an advertising message to their subsequent action, marketers realized that the Internet had leapfrogged over most other media forms to become the most accountable medium at their disposal (the exceptions being direct mail and direct response messages on TV or radio). Thus, the drive for accountability has helped speed up the development of digital metrics for other media. In TV, for example, the digital data coming off the cable set top box (STB) or from a digital video recorder (such as TiVo) provides a ‘clickstream’ of viewing activity at the household level that can report on second-by-second behavior in the household. From this, marketers can gain access to several new metrics. Included in these are:

- **Commercial ratings**: percent of households viewing an individual commercial
- **Commercial retention**: percent of households viewing the program who remained tuned during the commercial
- **Tuneaway**: among homes tuned to a channel prior to a commercial, the percent that tuned away at some point during that commercial
- **Timeshifting**: in Digital Video Recorder (DVR) households, the amount of time that is viewed at a time later than when the program or commercial first aired

Although for some in the industry, the return to household-based measures from the demographic (age and sex) viewer groups available in the syndicated ‘currency’ measures is considered a step backwards, others believe it is in fact a liberating move. There is increased skepticism that the traditional age/sex segments are in fact discriminating enough to show meaningful differences in the viewing behavior. Then, and perhaps more importantly, the household STB data can be readily matched up to more discrete target groups, including a marketer’s proprietary database of customers. This is the promise of addressable TV advertising, whereby different ad messages are sent simultaneously to different homes, based on the known characteristics of those homes. For example, dog food ads would only be sent to homes known to own a dog. Households with children could receive an ad for a minivan while the single young male living next door would, at that exact same moment, see an ad for a sports car. The data to create such matches come from third party companies such as Acxiom or Experian, that collect data on all households as part of their credit reporting services. Although, as of 2009, there are no nationwide technical standards, addressable advertising is being tested at regional levels (Helm 2009). Eventually, the ad viewing may be connected directly to purchase: consumers will be able to obtain information or
buy a product using their remote controls. It is also possible that the targeted devices could be used for copytesting several versions of an ad or for commercial delivery that is linked to prior ad exposure. If a home has seen one particular McDonald’s ad 20 times, it would then get switched, and a different piece of creative could be sent to that household, thereby reducing the likelihood that the target viewers would be tired of seeing it and tune it out.

Despite the advantages for creating targeted advertising, thereby eliminating advertising waste, the questions raised by digital media and metrics are many. In a world where one’s Internet or TV activity can be tracked down to the second, many fear a “big brother” scenario where someone (or some company) could use that information to ill effect. If a consumer is searching on the web for information about diabetes, would her insurance company find out and raise her premiums? Could a parent’s heavy viewership of “trashy” television be used against them in divorce proceedings? And does the fact that a person watched an ad for a particular car (or beer or shoe) mean that he is then going to be bombarded by other messages from that same marketer because they believe he is more interested in their offering?

What all of this comes down to is the issue of consumer privacy. Who owns the data that is tracking consumers’ media movements? Who should have access to that data? Can the data even be collected if one has not expressly and overtly given his or her permission in advance? And are there various levels of acceptability? That is, is it tolerable to collect and report aggregated household information but only after a set minimum of homes have been included? Once consumers opt in to a panel, is everything that they are doing reportable, or are there still limits on the data that can be collected? All of these issues are currently being debated and investigated. A focus on privacy issues is presented next.

**PRIVACY**

“The same technological advances that have made the Internet a potent marketing tool have also multiplied the threats to user privacy” (Lwin, Wirtz, & Williams 2007, p. 572).

In offline environments, consumers are generally willing to forego a certain amount of privacy to take advantage of benefits from marketers. For example, retail-shopping cards collect details about individuals’ shopping behaviors. In exchange for providing a certain amount of personal information, the consumer receives coupons or other incentives for products targeted to her. Even if privacy is a concern, consumers feel they can control the amount of information provided and weigh the benefits of doing so. However, technology has changed the scope and nature of information gathering (Ashworth & Free 2006). The scope of data available is seemingly infinite. For example, Yahoo! “collects 10 terabytes of user data a day, not including content, email or images”… (Conti & Sobiesk 2007, p.112). Although these data are used to segment and target messages more effectively to the consumer, the concern persists that information is gathered online in such a way that the audience cannot detect or avoid it (Ashworth & Free 2006). Further, as Milne (2000) points out, consumer information is stored “on a database platform that is potentially accessible to the entire Internet world” (p.1).

Clearly, the scope and methods of data gathering and storage can pose greater risk to consumers. For instance, in 2006 America Online (AOL) inadvertently released their dataset containing 20 million web searches for 657,426 of their members. The data included search queries—some containing potentially sensitive information related to medical or financial issues. Although the data were in fact anonymous (numbers were used, not names), it was apparently fairly easy to work backwards
from the “anonymous” cluster of web searches to identify real-world users (Barbaro & Zeller 2006). This privacy leakage was related to how the company stored the information, yet issues arise with all facets of digital metrics. Here, we review consumers’ view of privacy issues and then discuss the issues with respect to collecting, monitoring and learning from digital metrics. In so doing, we outline future directions and questions for research. See Figure 1 for a conceptual diagram.

**CONSUMER PRIVACY: KNOWLEDGE, CONTROL AND CONCERN**

Consumer privacy has been conceptualized in academic research as “the consumer’s ability to control (1) presence of other people in the environment during a market transaction or consumption behavior; and (2) dissemination of information related to or provided during such transactions or behaviors to those who were not present” (Goodwin 1991, p. 152). With digital media, other people may not be physically present, but the method of measurement can mean data are being collected, stored, and perhaps disseminated without a consumer’s knowledge or control.

Two important dimensions of privacy include consumer knowledge (high/low) and consumer control over information collection and use (high/low; Milne 2000). In order to exert some form of control over data, consumers should have some knowledge about data collection and use procedures in general—i.e., a “general online marketing literacy”—perhaps best captured by the Persuasion Knowledge Model (PKM) (e.g., Friestad & Wright 1994, 1995, 1999). The PKM examines the general set of beliefs that lay people hold about how persuasion “agents” (e.g., marketers, PR practitioners, advertisers) operate, including...
perceptions of agents’ goals and tactics, evaluation of the effectiveness or appropriateness of persuasion attempts, and self-reflections of agents’ own ability to cope with these attempts. It is essentially a ‘folk model’ of how persuasion attempts work. People may learn about persuasion tactics such as online advertising from many sources, including discussions with family and friends, direct experiences with persuasion agents, and commentary in the media (Friestad & Wright 1994). In addition, what people believe they know (perceived persuasion knowledge) and their actual knowledge may differ.

Research has shown that people believe that they learn about privacy issues from privacy seals and privacy notices, yet their ‘actual’ knowledge is very low (Rifon et al. 2005; Turow et al. 2008). It is believed that consumers are generally unaware of cookie practices or show overconfidence in their perceived knowledge as compared to real knowledge about such data collection procedures (Miyazaki 2008). For example, whereas 83% of respondents from a nationally representative U.S. sample correctly believed that “companies today have the ability to follow activity across many websites on the Web,” 25% incorrectly believed that “when a website has a privacy policy, it means the site will not share information with other Web sites or companies” (Turow et al. 2008, p.416). These findings led the researchers to conclude that only “a small proportion of Internet-using American adults have a highly sophisticated knowledge framework regarding marketplace privacy” (p.419). This survey offers some insight into persuasion knowledge about privacy—but a more detailed approach to understanding the level of persuasion knowledge about online advertising or metrics is warranted. What characteristics or tactic knowledge is essential for consumers? For example, what is the level of knowledge related to the terminology and functionality of various advertising forms, as implied by the deputy general counsel for Google, Nicole Wong’s question: “…would the user really understand what a behaviorally targeted ad is compared to a contextual ad?” (Hansell 2009).

There is also media literacy related to the potential risks and benefits associated with specific information collection (Rifon et al. 2005) and with privacy rules in general (Turow, Hennessy, & Bleakley 2008). When the online world is constantly shifting, how might everyday consumers gain persuasion literacy? Who has the responsibility for online literacy? Are there demographic or psychographic factors related to levels of media literacy on these issues? Future research might address these questions.

- **RQ 1:** What is general tactic “persuasion knowledge” in the context of digital media? (i.e., what exactly do consumers need to know?)
- **RQ 2:** What is the level of persuasion knowledge about online advertising, digital metrics, and behavioral targeting?

Research has shown that most people suspect their data are being used, but they don’t know how (Turow 2003). Further, despite the vast media coverage surrounding the AOL leakage, a majority of university students and middle-aged adults surveyed within six months of the incident were not aware of it (Conti & Sobiesk 2007). In addition to general online literacy, consumers should also have specific knowledge about particular company practices with regard to data collection and use, gained through experience and by reading companies’ privacy statements or other notices. The information can give consumers control—they can decide whether they wish to provide information or even visit the web site (Culnan & Milberg 1998). Although most companies now have online privacy statements (Adkinson, Eisenach, & Lenard 2002), a survey by TRUSTe’ revealed that only 20% of people read privacy statements “most of the time.” Reasons for not reading statements include perceived difficulty of understanding the notice (Good et al. 2005), distrust that the company will
Digital Metrics

adhere to policies, and a general apathy or lack of concern about privacy (Milne & Culnan 2004). To combat these issues, researchers advocate creating standard, readable notices or even opt-in policies on web sites (Milne 2000).

Opt-in policies would give the consumer the most control over collection and use of their information. Indeed, this position was advocated (but not mandated) by Federal Trade Commission Chairman Jon Leibowitz (Davis 2009b). However, some experts suggest there is no practical structure or format available in the ad industry as of 2009 (Kalehoff 2009). Instead, researchers advocate standardized “iconic representations” that offer easily recognizable and specific communication about privacy practices of the web site (Rifon et al. 2005; Turow et al. 2008). Yet, even readable notices—up front and not buried within a privacy notice—are uncommon. Also, given that any web site likely contains multiple advertisements with each one of them collecting, storing and sending different types of information about the viewers—it is difficult to create a standard notice (Hansell 2009). In response, however, Google instituted a policy for its behavioral tracking system whereby the phrase, “Ads by Google” is placed on each ad. If a user clicks on the phrase, she will receive some information about tracking and how to ‘turn off’ some of the tracking functions. Turow, privacy expert and professor at the Annenberg School for Communication, recommended a more comprehensive system (Hansell 2009). This system places the letter “T” for targeting on each ad to alert users that ad is collecting information about them. If users click on the ad, they are directed to a ‘privacy dashboard,’ where they could learn exactly which information was used to target the user and how to edit the information or opt out. In 2009, at least one web publisher experimented with providing specific knowledge to its readers (Kalehoff 2009). A banner ad on the “All Things Digital” blog web site of the Wall Street Journal included a notice that read, “A note about tracking cookies” (Kalehoff 2009). The notice went on to inform the viewer of the existence of tracking cookies on the web site and how to get rid of the tracking cookies. Further, if viewers clicked on links within the notice, they discovered additional educational materials about cookies.

Yet, research to date has not investigated the influence of these various labeling practices on consumer knowledge or behavior.

- **RQ 3:** How might general persuasion knowledge about online advertising and behavioral targeting influence self-efficacy and audience response?
- **RQ 4:** How might specific persuasion knowledge of a particular advertisement - gained through ‘notice’ - influence audience response?

Several studies suggest that as consumers gain knowledge of online procedures in general, they demonstrate more accurate knowledge about data collection practices (Turow et al. 2008) and are able to more critically evaluate and control their own behavioral choices (Good et al. 2007; Miyazaki 2008). Indeed, simple disclosure of specific cookie use to customers before they visited a web site significantly diffused their negative reactions (Miyazaki 2008).

Perhaps those with high knowledge and high control have a higher sense of **privacy self-efficacy** (Rifon et al. 2005). Self-efficacy is the “beliefs in one’s capability to organize and execute a particular course of action to achieve important attainments” (Bandura 1997, p. 3) and privacy self-efficacy is the confidence in one’s ability to protect one’s privacy. Those who score higher on privacy self-efficacy were found to provide less information on a website (Rifon et al. 2005). Yet, those with high knowledge and control are sometimes also willing to accept some personal data collection to benefit from incentives (Lwin, Wirtz, & Williams 2007) perhaps because they feel they can control the information.
**RQ 5:** How might privacy self-efficacy influence audience response to online advertising and privacy?

The response to privacy measures and issues in general is not standard across audiences. Older people and women have been shown to be more concerned than men about online information gathering and ramifications for personal privacy (Milne et al., 2004; Sheehan 1999). In a survey of younger people (university students), the concern over digital privacy was not prevalent (Conti & Sobiesk 2007). Users felt they had adequate knowledge and control over their own information—and the majority felt it was their own responsibility to protect privacy. In other studies, years of schooling and those with more online experience were more likely to generate more accurate beliefs about online data collection (Turow et al. 2008) and to report that they engage in ‘protective behavior’ (Milne et al. 2004).

Beyond demographic variables, there is a standard Privacy Segmentation Index (PSI) created by market-research firm, Harris Interactive, which queries people about their attitudes toward privacy and then categorizes them into three “privacy-sensitive groups”: (1) privacy fundamentalists—who feel strongly about privacy; (2) privacy pragmatists (who have strong feelings over privacy related to information misuse, yet allow information collection if reasons for use are provided); and (3) “privacy unconcerned”—those with no real concerns (Taylor 2003). Yet, the implications of these individual-difference variables for privacy behaviors have not been fully explored.

**RQ 6:** How do individual differences in attitudes toward privacy (or privacy concern) influence subsequent behaviors?

Some research has examined relationships between attitudes or concerns over privacy and subsequent behaviors (e.g., Rifon et al. 2005; Celsi & Olson 1988). Yet this literature has revealed mixed results in part due to the problems with inconsistent terminology and measurement of the construct “concern for privacy” and the potential for social desirability in responses (Rifon et al. 2005). In one such study, university students claimed to perform various behaviors to protect their online identity such as encrypt email, use anonymous re-mailers, and use anonymizers while browsing, yet their responses also correlated positively with a social desirability index (Milne et al. 2004). The authors comment that “students might have overstated their technical abilities” (p.225). Indeed, no perceived or actual knowledge was assessed in the study.

Celsi and Olson (1988) suggest that greater concern for privacy might motivate information processing and behaviors related to privacy, such as seeking out privacy related information. Other studies have also shown links between attitudes or concern over privacy and self-reported behaviors. For instance, in an Internet survey, Sheehan and Hoy (1999) noted that concerns over privacy were related to users’ subsequent behaviors—such as not registering for a web site, providing incomplete information, requesting removal from a mailing list and even ‘flaming’ (sending highly negative responses). Similarly among Australian and South African Internet users, Dolnicar and Jordan (2007) found that privacy concerns exist and that such concerns are related to self-reported ‘protective’ behavior. Despite these findings, other researchers suggest there may be no such relationships (Rifon et al. 2005). What is needed is a standard definition of the key concepts (e.g., privacy concern) and empirical research to assess behavioral intentions and actual behaviors. For example, would concern over privacy influence whether or to what extent consumers choose to interact with digital advertising?

**RQ 7:** How does concern for privacy influence consumer behaviors and responses to online advertising?
Concerns over data collection usually fall into one or more of the following categories; collection of personal information, internal and external unauthorized use of secondary use of information, error in personal information, and improper use of information (Smith et al. 1996). Another taxonomy (Wang et al. 1998) shows that people are worried about improper data acquisition (access, collection or monitoring), improper use (analysis or transfer), privacy invasion (unwanted solicitation), and improper storage. Indeed, data retention or storage is rarely discussed, but “anecdotal evidence suggests that every interaction with these companies is scrupulously logged and stored indefinitely” (Conti & Sobiesk 2007, p.112). The database is one of the most valuable assets of online companies. Despite the potential for leakage of the AOL variety, a survey of university students revealed little concern over data storage (Conti & Sobiesk 2007). Indeed, 94% of respondents indicated that “indefinite search retention” would not change their surf habits. The issue with such retention relates to company changes in the future or loss of control of that valued database.

Advertisers and consumers alike need to understand privacy issues related to methods of data collection and use. One of the most popular social network web sites, Facebook, has faced a number of privacy issues with respect to data collection and use. Thus, it serves as an apt case study to analyze consumer privacy concern and policies.

**CASE STUDY: FACEBOOK**

**Beacon Blunder and Beyond**

Facebook is a social networking web site that was founded by Mark Zuckerberg and friends at Harvard University in 2004. What began as a pilot project for use by Harvard students expanded to high school and university students and finally to anyone with a valid email address. As of 2009, the web site boasts a user base of 200,000,000 active users worldwide. Although the primary function may be for users to connect to one another and share information, its use as a marketing device through the use of banner ads and other products and services such as ‘branded groups’ (e.g., Apple) is noteworthy. Yet, in its relatively short lifetime, the web site has faced a number of challenges to privacy.

An early Facebook service “Techmeme” was based on the power of “word of mouth” and friend referrals. The service was essentially a ‘social ad,’ which worked by broadcasting to users the name and photo of their friends who like or used a certain brand, product, or service (Hansell 2008). The service, launched in November 2007, used to discern which products users like was called Beacon, which gathered information about what you like (or bought) from other Beacon-affiliated web sites. According to a company press release, the service would allow users to perform various actions related to the participating web sites, including “posting an item for sale, completing a purchase, scoring a high score in an online game or viewing of video.” Although the idea of sharing branded information with friends related well to the overall use of Facebook, the problem occurred largely because the service was initially opt-out, and users’ friends received updates and status feeds, sometimes related to their own gifts! As a result, advocacy group Moveon.org organized a petition and facebook group with the following message just a couple weeks after the Beacon launch:

“Matt in New York already knows what his girlfriend got him for Christmas...Why? Because a new Facebook feature automatically shares books, movies, or gifts you buy online with everyone you know on Facebook. Without your consent, it pops up in your News Feed—a huge invasion of privacy. Can you sign the petition to facebook today? Then invite friends to this group! Petition: “Facebook must respect my privacy. They should not tell my friends what I buy on other sites—or let companies use my name to endorse their products—without
my explicit permission.” Apparently, within 10 days, more than 50,000 people joined this group and as a result, Facebook changed the service from opt-out to an opt-in service (Farber 2007) with a public apology and new privacy controls announced by Mark Zuckerberg in December 2007 (Zuckerberg 2007).

In addition to privacy issues surfacing from Beacon, Facebook has come under scrutiny for the ability of hackers to download users’ profiles through data mining (Jones & Soltren 2005) and when users were unable to completely remove their profiles (Aspan 2008). At stake was the ownership of user data—would Facebook retain the information that users uploaded even after the users deleted their profiles? According to a terms of service agreement, it appeared that Facebook owned users’ data ‘indefinitely.’ After this policy came to the attention of the media and Facebook users by an article in The Consumerist (Walters 2009), Facebook agreed to change its terms of service agreement. As a result of these various privacy issues, in February 2009, Facebook involved its users in the creation of policy and principles of its web site. According to the company’s press release: “Users will have the opportunity to review, comment and vote on these documents over the coming weeks and, if they are approved, other future policy changes.” In essence, the popular web site evolved from publisher-centered information ownership and control to a more consumer-centric and democratic model of data ownership control and policy. Despite the growing popularity of the web site, questions surrounding its value and future related to advertising revenue are still raised. The future of Facebook and the monetization of social networking and advertising remains a mystery.

PRIVACY SUMMARY

Digital metrics offer an enormous amount of information about “the audience”—explicit information provided willingly by the consumer through web site registrations, profiles (such as on Facebook) as well as implicit information collected by cookies. Clearly, the lines between what digital measurement can provide and what consumers are willing to have collected about them are still being clarified. There will be more consumer education needed, to explain the benefits and value to the consumer of passive digital measurement, and the possible threats that a loss of personal data can create. Turow et al. 2008 suggest, for example, that without a unified national philosophy of marketplace privacy—“the best approach for educating Americans on the subject may well be to streamline the discussion of regulations”—especially through schools, community organizations, and the media (p. 421). Marketers will keep focused on the former, while privacy advocates and consumer groups will likely emphasize the latter. It is in the industry’s interests to educate consumers about online data and how it is and should be used. Websites should all post their privacy statements in clear and accessible ways, including the presence and use of cookies. For the inclusion of a discrete amount of useful information at the time when the data are actually being collected would help consumers’ knowledge acquisition and control (Cranor 2003).

FUTURE RESEARCH AND CONCLUDING THOUGHTS

There has never been a more exciting time to be dealing with audience measurement and media usage. Consumers work and play with digital media, advertisers are increasingly turning to digital forms for their messages, and regulators are considering the ethical and legal ramifications of the digital media. Yet, academic research has not yet fully embraced research in digital media. We hope that this chapter will serve as a catalyst to move research forward in this area.

The growth and development of digital media forms have provided consumers with more
choice and control over their media consumption, at the same time as they offer marketers more accurate and granular information on which to base their spending decisions. We have outlined some common digital metrics used today, but what will the future of digital media look like? Will advertising remain a key revenue model for websites? In 2009, some advertisers are experimenting with new 3D forms. For example, Carmichael Lynch created a 3-D Sasquatch character for Jack Link’s that users could print out and play with in front of a webcam to interact with a 3D digital environment (Klaassen 2009). Marketers are also increasingly interweaving marketing material with editorial content (Hessel 2009). For example, The New York Times offered a web advertising campaign for the television series, Mad Men, which included a “mini-archive of Times articles about the show within the ad unit.” Said the Vice President of Research & Development, Michael Zimbalist, “We have to give advertisers an opportunity to market through the content, not just around it.” Further, as mobile technology catches up, digital media has finally moved beyond the laptop into mobile opportunities where geo-targeting is not uncommon. Each of these new forms offers new avenues for exploring ad effectiveness and social issues. Are the ads more likely to be accessed if integrated into content? If targeting the consumer where she lives and plays? What has happened to the so-called Chinese Wall separating editorial and advertising in a digital age? Questions such as these offer promising areas for future consideration in academic research and professional practice.

In fact, digital media are changing all media. By offering marketers greater accountability for their media dollars, these new media forms have set a new standard for marketers’ understanding of and expectations for the role that all media play in the marketing mix. As the research director at one large media agency put it, “There is more and more emphasis by advertisers for greater return-on-objects in campaigns, particularly in the digital space where the accountability data is so readily available” (New Study Shows…2008). Indeed, as Jack Klues, president of Publicis Groupe, has noted, marketers increasingly will want and need to begin the media planning process with digital (online) media, rather than adding them in as an afterthought to a television-based plan [Klues 2008].

If this is, indeed, the path that marketers take, the need for even more digitally-based audience measurement will grow. There will be continued experimentation, for example, with multi-screen measures (TV + PC + Mobile) now being tested by companies such as Nielsen, as well as continued work in data fusion. Media data will increasingly be combined with purchase information to provide a more holistic view of how all the elements truly work together to achieve a marketer’s ultimate objective, to increase sales. At the same time, we will inevitably see the development of industry standards to ensure that the measurement services continue to provide reliable and valid data, and do so in ways that do not infringe on consumers’ rights in any way. The debate over privacy is bound to get louder and will likely involve legislators and regulators along the way. For example, Representative Rick Boucher (D-VA.) publicly stated he will introduce privacy legislation in 2009—while FTC Chairman Leibowitz appears to side with consumers’ rights (advocating an opt-in policy). Yet, at the highest level, Supreme Court Justice Antonin Scalia warns that privacy laws may conflict with the First Amendment (Davis 2009a).

In the end, what we hope and expect the development of digital metrics to do is give John Wanamaker, and all the marketers that followed him, a better answer to the question of “which half” is wasted. And in doing so, help them determine ways to avoid that waste altogether.
REFERENCES


Klues, J. (2008, April 11). Media Today. Presentation to University of Illinois at Urbana-Champaign Department of Advertising Faculty, Illini Center, Chicago.


