Journal of Broadcasting & Electronic Media/March 2002

The Internet Audience: Web Use as Mass Behavior

James G. Webster and Shu-Fang Lin

We conceive of the Internet as a medium of mass communication, and we analyze 2 behavioral features of its audience: size and duplication. Internet audiences are highly concentrated in the most popular Web sites, conforming to Pareto's Law. Duplication across pairs of Web sites is explained by the size of the audiences unique to each site in the pair, conforming to Goodhardt's duplication of viewing law. In addition, similarities of content or domain modestly increase duplication across Web sites. A 4-variable model explains over 80% of the variance in duplication.

At last count, over 160 million people in the United States had access to the Internet, and in a typical week nearly 100 million actually logged on (Nielsen// NetRatings, 2001). Whether e-mail, e-commerce, or just "surfing," this activity is, often times, a matter of individuals visiting specific sites on the World Wide Web (hereafter the Web). Whatever else it is, Web use can be thought of as a kind of mass behavior, similar in many respects to choosing TV programs or publications. Indeed, it has been argued that the Web can be seen as a medium of mass communication and its users as a mass audience (Morris & Ogan, 1996; Roscoe, 1999). Studying Web use in the aggregate, as we do here, reveals certain law-like regularities not unlike those found in more traditional mass media. Knowing these patterns can enhance our understanding of this new medium and its potential to affect society.

We use the term "mass" with some trepidation. To many in media studies, the word conjures up the specter of "mass society" with its connotations of passivity, susceptibility to influence, and vulgar tastes (see Beniger, 1987; Williams, 1958/ 1983). We make no such claims one way or another, but use the term "mass behavior" in a more limited sense. Following Blumer (1946), we conceive of a mass as a large, heterogeneous collection of individuals who act autonomously and are, for the most part, anonymous. This is a well-established way of conceptualizing

James G. Webster *(Ph.D., Indiana University) is a Professor in the Department of Communication Studies and Senior Associate Dean in the School of Speech at Northwestern University. His research interests include audience studies and the impact of new media.*

Shu-Fang Lin *(M.A., Northwestern University) is a Ph.D. student in the School of Journalism and Communication at the Ohio State University. Her research interests include audience studies and the uses and effects of mass media.*

The authors thank Nielsen//NetRatings for making the necessary data available for analysis. The results are based on Lin's unpublished master's thesis at Northwestern University.

 Journal of Broadcasting & Electronic Media 46(1), 2002, pp. 1-12

social collectives (McPhee, 1963; Porter, 1986; Webster & Phalen, 1997), and it seems nicely suited to Web users. Moreover, by applying this analytical model to the Internet, we bring that behavior within the ambit of research and theory on mass audiences.

Specifically, our approach is to take Internet Web sites and their attendant audiences as the unit of analysis. We focus on two important features of mass behavior associated with Web sites: audience size and audience duplication. Audience size is critical to virtually all forms of subscriber or advertiser-supported media. Electronic and print media depend upon audience "ratings" and circulation to sustain their operation. The Internet is no exception. According to an industry trade organization, in 1999 over $4 billion was spent on Internet advertising (Internet Advertising Bureau, 2000). The sheer size of audiences can also be an index of a medium's cultural significance or its potential to cultivate various effects (Dayan & Katz, 1992; Webster & Phalen, 1997). Audience duplication is a form of cumulative behavior that underlies such things as frequency of exposure, audience flow, and audience loyalty. These are of relevance to both advertisers and programmers (Webster, Phalen, & Lichty, 2000), and they reveal something about longer-term patterns of exposure to the Internet and the intensity with which people use different types of Web sites.

Audience Size

There are two strains of literature that deal with audience size. The first concentrates on the determinants of audience size or market share. Such factors include audience needs or preferences, the amount of promotion associated with media products, the seasonal or temporal availability of audiences, and a host of structural features within the media themselves that induce people to watch, read, or tune in to a particular item of media content (Webster & Phalen, 1997). The second takes audience size as it finds it and makes something of how those audiences are distributed. Work on popular culture, cultivation, media economics, and audience fragmentation is of this sort. Here the sheer magnitude of the audience, or lack thereof, is of consequence. It is the latter strain of literature that we will draw on in this article.

In particular, we consider the extent to which Web site audiences show evidence of what is sometimes called "Pareto's Law." Vilfredo Pareto (1848-1923), an economist and political sociologist, "discovered" that a small portion of a nation's population accounted for a disproportionate amount of its income. This markedly asymmetric pattern was evident throughout history in countries with very different economic systems, leading Pareto to assert an empirically determined law of income distribution (Persky, 1992). More recently, marketing researchers have come to recognize a similar pattern in the consumption of consumer goods. This is sometimes called the "80-20 Rule," which states that 80% of a brand's sales volume is accounted for by 20% of its buyers (Anschuetz, 1997). While claims of such law-like

behavior have been subject to criticism, these patterns persist in many different contexts.

The audiences for traditional forms of mass media show evidence of this lopsided distribution. For example, 10% of the movies released in any given year account for roughly half the box office revenue, conforming to a Pareto distribution (De Vany & Walls, 1999; Neuman, 1991; Vogel, 1998). There are thousands of book publishers in the United States, yet the top 10 account for over 60% of all book sales (Compaine & Gomery, 2000). Graphically, these data have been represented as an upward sloping curve, called a "Lorenz curve" (see Neuman, 1991). Most cultural products, be they television programs, magazines, or records, show the same kind of asymmetrical distributions wherein the vast majority of viewers, readers, and listeners are concentrated in the top-rated items of each category (Ehrenberg, Goodhardt, & Barwise, 1990; McPhee, 1963).

Does the Internet show similar signs of asymmetry? Rather than being concentrated—there are several million Web sites available to Internet users—mass audiences might be dispersed in ways unparalleled in traditional forms of media. Preliminary analyses reported by researchers at Xerox found that among AOL users "the distribution of visitors per site follows a universal power law similar to that found by Pareto in income distributions" (Adamic & Huberman, 1999, p. 3). Specifically, their data show that the top 5% of sites accounted for almost 75% of user volume. These results, however, were based only on AOL subscribers and may therefore lack generalizability. We examine similar patterns in the distribution of Web site audiences, based on a broader sample of Internet users.

Audience Duplication

The literature on audience duplication is more specialized and typically confined to marketing and media studies. Audience duplication deals with the extent to which the people exposed to one item of media content (e.g., a magazine or television program) are exposed to other such items. Beginning in the early 1960s, mathematical models of audience duplication were tested in different forms of mass media (e.g., Agostini, 1961; Bower, 1983; Cannon, 1983; Kuhn, 1983; Leckenby & Hong, 1998; Rust, 1986). Television audience behavior was the most closely scrutinized (e.g., Goodhardt, Ehrenberg, & Collins, 1987; Headen, Klompmaker, & Rust, 1979; Webster & Phalen, 1997). Perhaps the best known statement of the law-like behavior of television audiences is Goodhardt's "duplication of viewing law" (Goodhardt, 1966; Goodhardt et al., 1987). It states that the audience common to any pair of television programs (i.e., the duplicated audience) is a function of the rating of the first program multiplied by the rating of the second program times an empirically determined constant. For two programs on the same network, the constant is usually in the neighborhood of 1.6, demonstrating a kind of channel effect called "channel loyalty." Subsequent work has refined this model by introducing other factors that

help explain levels of audience duplication (e.g. Headen et al., 1979; Henriksen, 1985; Webster, 1985), but the essential features of the law remain unchanged.

To unbundle these mathematical models a bit, three broad categories of factors have been considered when explaining variation in the size of the audience common to any pair of items. The first and most powerful determinants are the audience sizes for each item (e.g., a program, a magazine, etc.) in the pair. For example, if no one watches a television program (i.e., its rating is 0), it will have no audience in common with any other program. Conversely, if everyone watches a program (i.e., its rating is 100), it will be perfectly duplicated with all other shows. Of course, such extreme ratings are rarely if ever seen, but the size of any one media audience determines in an inescapable way the possibilities for audience duplication. It is this statistical "fact of life" that is at the heart of the duplication of viewing law. By extension, to predict the audience duplicated across any pair of Web sites, one should begin with the size of each site's audience. These are the first two variables used in our study of duplication.

The second category to be considered is the content of the media involved. Several bodies of literature suggest that individuals have systematic preferences for content of a type. To the extent that people have enduring needs that can be gratified by certain content, they may seek out media of that sort (e.g., Rosengren, Wenner, & Palmgreen, 1985; Zillmann, 2000). Economic models of program choice posit even more explicitly the existence of program types "defined in terms of viewer preference" (Owen & Wildman, 1992, p. 72), meaning that people who choose one program of a type will tend to choose others of the same type. This theoretical supposition is buttressed by a number of empirical studies designed to discover viewer-defined program types (e.g., Kirsch & Banks, 1962; Rao, 1975; Rust, Kamakura, & Alpert, 1992). The consistent, if unremarkable, result of such work is that commonsense industry categorization schemes bear a modest systematic relationship to preference and/or choice. In the case of Internet users, it might be assumed that fans of a particular type of Web site (e.g. news, entertainment, search engines, etc.) move readily between sites within the category. Conversely, it might be that users who have learned to negotiate one site would tend to ignore others of the same type. In any case, it is reasonable to look for a "content effect" in audience duplication on the Web.

The third kind of factor that affects audience duplication is how items of media content are structurally related to one another. In traditional forms of electronic media, there is a wealth of research that demonstrates the impact of structure on patterns of duplication. These include channel loyalty (Bruno, 1973; Goodhardt et al., 1987), inheritance effects (Goodhardt et al., 1987; Webster, 1985), and repeat viewing (Ehrenberg & Wakshlag, 1987; Webster & Wang, 1992). Of course, the Internet does not have the same structural characteristics as radio and television, but it does have latent structures. Perhaps the most important are domains. These are families of Web sites, often under common ownership. Because members of the family are typically linked to one another, it seems likely that movement within

domains is slightly easier than movement across domains. This suggests an enhanced level of duplication between sites in the domain, directly analogous to channel loyalty. This study assesses that kind of "domain effect."

To summarize, our analysis of audience duplication on the Web considered four possible determinants. Factors one and two are the audiences unique to each Web site in the pairs we examine. For convenience, we distinguished between the larger Web site and the smaller Web site within each pair. The third factor was whether the sites in question were of the same or different types. The final factor was whether the sites were of the same or different domains.

Method

There were two parts to the analysis. The first examined audience sizes by arranging them in a Lorenz curve. The second analyzed audience duplication by applying a stepwise multiple regression procedure to selected pairs of Web sites. In both cases, the data were drawn from the Nielsen//NetRatings online audience report for the month of September 1999. This is a proprietary database made available to the authors for the purpose of this study. Nielsen//NetRatings produces audience estimates with a panel based on a probability sample of U.S. households with access to the Internet. Homes that agree to participate load software onto their computers that records Web usage. This is a "user-centric" form of measurement analogous to the peoplemeter system used by Nielsen Media Research to measure national television audiences (Webster et al., 2000). At the time the data were collected, the panel included 33,000 people. For a more complete description of methodology, see the Nielsen//NetRatings homepage (http://www.netratings.com/).

The analysis of audience duplication is the more complicated of the two and deserves a few words of explanation. Similar to studies of program audience duplication (e.g., Goodhardt et al., 1987; Headen et al., 1979; Webster, 1985), the unit of analysis was a pair of Web sites. Technically, this is referred to as pairwise duplication (Rust, 1986). Pairs were selected using a quota sampling technique. As Adamic and Huberman (1999) and our own analysis indicate, the top sites account for a disproportionate percentage of the traffic on the Web. To avoid problems of estimation attendant to small sites, we limited our universe to the top 200 sites as reported by Nielsen//NetRatings. Within that population, pairs were picked to represent one of four conditions formed by a 2 x 2 matrix. The first factor was the content dimension, coded as the same or different; the second factor was the structure dimension, coded as the same or different domain. A quota of 20 pairs was assigned to each condition, yielding a total of 80 cases for the multiple regression.

For each pair, we determined the "unique audience" of each Web site using the online report. This term is commonly used by Internet measurement companies and can be thought of as the reach, cume, or unduplicated audience for the site during

the month in question. This information was used to determine the relative size of each site in the pair, thereby labeling each as "bigger" or "smaller."

We adopted Nielsen//NetRatings' scheme for categorizing content. The top four categories were picked proportionally from our universe of 200 sites: search engines/ portals, entertainment, news and information, and computer hardware and software. Other content types were omitted because there were too few representatives to be analyzed.

The dependent variable in the regression (i.e., the duplicated audience) was determined by submitting each pair of Web sites as input to a tool on the Nielsen// NetRatings online report that estimated the absolute size of the audience common to any pair of Web sites. The four independent variables described above were used in a stepwise procedure to examine their relative ability to explain variation in audience duplication. All audience size variables were transformed logarithmically. This procedure is commonly used with similar data sets as a corrective for those variables not normally distributed (see Headen et al., 1979; Henriksen, 1985; Webster et al., 2000).

Results

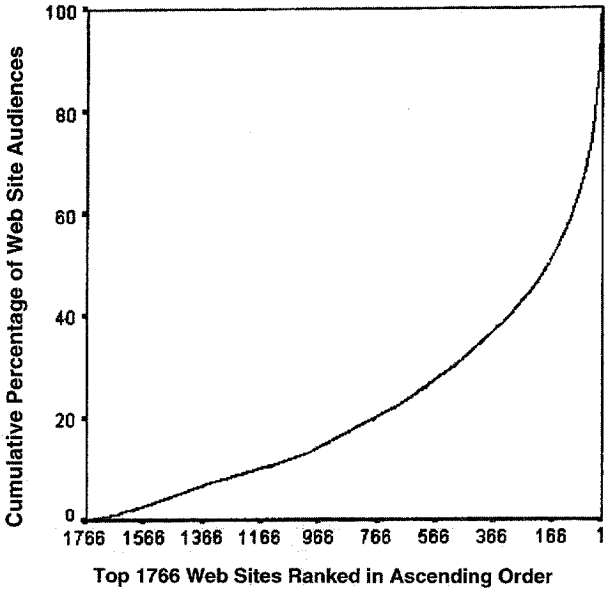
The Concentration of Audience Sizes

Nielsen//NetRatings reported the top 1,766 Web sites, rank ordered by their unduplicated monthly audiences. These ranged in size from almost 20 million visitors in a month (Yahoo!) to just fewer than 175,000 over the same time period. Casual inspection of the list, however, indicates that audience sizes drop off quite rapidly. The top 200 sites we used in the duplication study accounted for roughly half of all traffic on these Web sites. Among the full 1,766 sites for which we had data, the top 5% accounted for 37% of total visitors. The top 20% of sites accounted for 61% of total visitors. Of course, there are millions of Web sites available to Internet users. If these were figured into the analysis, we suspect we would, like Adamic and Huberman (1999), easily surpass the 80-20 rule-of-thumb.

These same data can be represented graphically. The most common visual description takes the form of a Lorenz curve, often used to portray the asymmetric distribution of incomes. As Neuman (1991) and others have done, we arranged our units (i.e., Web sites) in ascending order on the horizontal axis. Considering each unique audience as additive, we then plotted the contribution of successive units to the cumulative percent of audiences. If each Web site had an equally large unique audience, the Lorenz curve would, in fact, be a straight line rising at a 45-degree angle. This is sometimes called the "equality line." To the extent that the actual curve bows down from the equality line, there is evidence of concentration. In our case, as in other Pareto-like distributions, the curve ascends slowly across the majority of

smaller Web sites but turns sharply upward as the effect of the most popular sites is added in.

Figure 1
Concentration of Web Site Audiences



The Nature of Audience Duplication

As expected, the absolute size of the audience common to any pair of Web sites was, first and foremost, a function of the size of the audience unique to each site in the pair. The size of the larger Web site was highly correlated with the duplicated audience ($r = .765$, $p < .01$), as was the size of the smaller Web site ($r = .720$, $p < .01$). The majority of people who visited the larger site in a pair usually didn't visit the smaller site. However, of those people who visited the smaller site, almost two thirds had also visited the larger site. Since the larger sites tended to be search engines, this is perhaps unsurprising. On average, the level of audience duplication among the top 200 Web sites was 37% (i.e., just over one third of the people who visited any one site also visited the second site).

The other factors had a more modest role to play in duplication. The content of the sites involved was positively related to the levels of duplication ($r = .210$, $p < .05$), indicating that people who visited one site tended to visit others of the same general sort. The "domain effect" fell just short of statistical significance, though it was in the expected direction ($r = .176$). Table 1 shows the outcome of the stepwise regression

employing all four predictor variables. In combination, the audience size variables alone explain 76.1% (R^2 = .761) of the variance in Web duplication. Though less powerful determinants of duplication, the content and domain factors each added significantly to the predictive power of the equation, ultimately explaining over 80% of the variance.

Table 1
Determinants of Web Audience Duplication[a]

Independent Variables	Standardized Coefficients (t value)[b]			
	Step 1	Step 2	Step 3	Step 4
Audience size of big Web site (n. log)	.765	.555	.544	.606
	(10.485)	(9.071)	(9.430)	(10.607)
Audience size of small Web site	—	.476	.490	.409
(n. log)		(7.773)	(8.585)	(6.875)
Content	—	—	.185	.180
			(3.614)	(3.711)
Domain	—	—	—	.173
				(3.218)
Constant	−4.145	−5.582	−5.689	−5.882
R^2 (adjusted)	.580	.761	.794	.816
Overall F[c]	109.938	127.045	102.318	88.775
Df	1,78	2,77	3,76	4,75

[a]Dependent variable is in the form of the natural log of the duplicated audience of big and small Web sites.
[b]All figures in parentheses are t values significant at $p < .01$
[c]Each overall F value was significant at $p < .001$

Conclusion

The Internet supports different forms of communication, from the most private exchanges to the most public pronouncements. This study was a macro-level analysis of data on exposure to Web sites. Ignored are the needs and gratifications that motivate Web use, the disparate ways in which people employ Web sites, and the meanings they attribute to their encounters. Learned are the broad outlines of the Internet audience and how its patterns of exposure compare to those of more traditional media (e.g. movies, books, television programs, etc.). Specifically, 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. Internet audiences are concentrated in a relative handful of sites. As with many other cultural products that conform to Pareto's Law, attendance is dominated by the most popular offerings

(Neuman, 1991; McPhee, 1963). Similarly, patterns of audience duplication echo those found in Goodhardt's "duplication of viewing law" (Goodhardt et al., 1987; Headen et al., 1979).

These findings should be considered preliminary for a number of reasons that merit discussion. The Nielsen//NetRatings panel for September 1999 was composed entirely of home users of the Internet. It is conceivable that business users, or people visiting Web sites at different times of the year, exhibit different behaviors. Further, our use of broad monthly estimates of unique or common audience may mask many more nuanced features of individual behavior, including the time spent on sites or repeat visits to sites during the month.

Ideally, subsequent research should expand and randomize the sample of Web sites used to study Web site duplication. Though the top 200 sites alone account for roughly half of the traffic on the Web, including smaller sites would help capture the rich menu of offerings served by the Internet. Moreover, a random sampling of sites would better represent the population of Web site pairs as they fall into our 2×2 matrix. Of course, there are practical limits to either of these strategies. It is hard to reliably estimate the audiences for very small Web sites with panel data, and including all possible pairings creates an unimaginably vast number of pairs to be analyzed and coded.

The operationalization of both content and structural factors should also be considered. Our content categories were very broad. More discriminating categories are probably in order. Further, it is likely that different types of sites are used in different ways. For instance, search engines are among the most popular sites and probably serve a unique function as a gateway to other smaller sites. One might analyze duplication among only search engines or, for that matter, any other category of sites, looking for the impact of more subtle content distinctions within the group. The "domain effect" could also be examined more closely. We have supposed that sites within a domain are more frequently linked. Examining the actual structure of linkages—and their effect on movement across sites—could be illuminating.

Preliminary though they are, the results provide food for thought. With its millions of sites and services, it is easy to think of the World Wide Web as a thoroughly demassified medium (Brown & Duguid, 2000). It is not. Even the smallest sites we examined had 175,000 different visitors in the course of a month. Such large groups have the essential characteristics of a "mass." The people using sites are largely anonymous to one another and almost certainly act autonomously. Whether Internet audiences can be considered heterogeneous is a matter for further investigation. While some sites may, like other specialized media, appeal to a more targeted audience, others are likely to have broad appeal. The question of whether any given audience is homogeneous depends upon the attributes the researcher has chosen to examine. At any rate, it is doubtful that homogeneity would affect the behavioral characteristics associated with mass audiences.

It is also clear that not all Web site audiences are created equal. Internet audiences

are highly concentrated in a relatively small number of sites. This too argues against a notion of demassification in which many sites share audiences more or less equally. Instead, we find a Pareto-like distribution of audiences. Even in a market where barriers to entry are low, the fact that the Internet is dominated by a few large "speakers" has potentially troubling implications for public discourse (Netanel, 2000). Exactly why cultural products or incomes assume an asymmetric form has puzzled scholars since Pareto first reported his findings in the late nineteenth century. At the very least, we should be more aware of these disparities in a seemingly abundant medium and more cognizant of their economic and social implications.

The results should also assuage fears of two unpleasant social consequences that could conceivably flow from increased Web use. The first has to do with the public's diet of Internet content. Because the Web makes rather specialized content available around the clock, some have imagined a kind of "tunnel vision" in patterns of consumption (Brown & Duguid, 2000). If individual users seek only a limited menu of those things that gratify their needs while ignoring all else, a kind of social polarization might result (Webster & Phalen, 1997). The lack of a pronounced "content effect" in our study of audience duplication suggests that people opt for a more varied diet than the worst fears would suggest. As the content categories we used were very broad, a more refined analysis would be in order. Still, the initial findings are hopeful.

A second troubling prospect for Internet users stems from the understandable desire of content providers to manage the flow of visitors to their Web sites (Lessig, 1999). Might people looking for the easiest way to navigate the Web be unwitting prisoners of a family of sites? Domains are an important, if sometimes invisible, organizing structure for Internet users. Yet the absence of a clear "domain effect" suggests that people are not easily corralled within a single domain but move freely from one domain to another. Here again, our results are preliminary. Since portals and search engines are among the most popular sites, it would be especially instructive to analyze the flow of traffic to and from this particular kind of Web site.

Some years ago, Raymond Williams wrote "There are in fact no masses; there are only ways of seeing people as masses." (1958/1983, p. 300). Certainly the Web, with its diverse set of offerings, invites a kind of micro-level analysis that addresses each person's experience with the Internet (e.g. Lindlof & Shatzer, 1998). There is value, however, in seeing people as a mass. Doing so allows researchers to apply well-established bodies of research and theory to address macro level questions of social use and impact.

References

Adamic, L. A., & Huberman, B. A. (1999). *The nature of markets in the World Wide Web.* Palo Alto, CA: Xerox Palo Alto Research Center.

Agostini, J. M. (1961). How to estimate unduplicated audiences. *Journal of Advertising Research, 1,* 11-14.

Anschuetz, N. (1997). Profiting from the "80-20 rule of thumb." *Journal of Advertising Research, 37,* 51-56.

Beniger, J. R. (1987). Toward an old new paradigm: The half-century flirtation with mass society. *Public Opinion Quarterly, 51*(Suppl.), S46-S66.

Blumer, H. (1946). The field of collective behavior. In A. M. Lee (Ed.), *New outline of the principles of sociology* (pp.167-222). New York: Barnes & Noble.

Bower, J. (1983). Net audience of U.S. and Canadian magazines: Seven tests of Agotini's formula. *Journal of Advertising Research, 23,* 13-20.

Brown, J. S., & Duguid, P. (2000). *The social life of information.* Boston: Harvard Business School Press.

Bruno, A. V. (1973). The network factor in TV viewing. *Journal of Advertising Research, 13,* 33-39.

Cannon, H. M. (1983). Reach and frequency estimation for specialized target markets. *Journal of Advertising Research, 23,* 45-50.

Compaine, B. M., & Gomery, D. (2000). *Who owns the media? Competition and concentration in the mass media industry* (3rd ed.). Mahwah, NJ: Erlbaum.

Dayan, D., & Katz, E. (1992). *Media events: The live broadcasting of history.* Cambridge, MA: Harvard University Press.

De Vany, A. S., & Walls, W. D. (1999). Uncertainty in the movie industry: Does star power reduce the terror of the box office? *Journal of Cultural Economics, 23,* 285-318.

Ehrenberg, A. S. C., Goodhardt, G. J., & Barwise, T. P. (1990). Double jeopardy revisited. *Journal of Marketing, 54,* 82-91.

Ehrenberg, A. S. C., & Wakshlag, J. (1987). Repeat-viewing with people meters. *Journal of Advertising Research, 27,* 9-13.

Goodhardt, G. J. (1966). The constant in duplicated television viewing between and within channels. *Nature, 212,* 1616.

Goodhardt, G. J., Ehrenberg, A. S. C., & Collins, M. A. (1987). *The television audience: Patterns of viewing (2nd ed.).* Westmead, UK: Gower.

Headen, R. S., Klompmaker, J. E., & Rust, R. T. (1979). The duplication of viewing law and television media schedule evaluation. *Journal of Marketing Research, 16,* 333-340.

Henriksen, F. (1985). A new model of the duplication of television viewing : A behaviorist approach. *Journal of Broadcasting & Electronic Media, 29*(2), 135-145.

Internet Advertising Bureau. (2000, April 18). *Internet advertising revenues soar to $4.6 billion in 1999.* Retrieved from (http://www.iab.net/).

Kirsch, A. D., & Banks, S. (1962). Program types defined by factor analysis. *Journal of Advertising Research, 2,* 29-31.

Kuhn, W. (1983). Net audience of German magazines: A new formula. *Journal of Advertising Research, 23,* 30-33.

Leckenby, J. D., & Hong, J. (1998). Using reach/frequency for Web media planning. *Journal of Advertising Research, 38,* 7-20.

Lessig, L. (1999). *Code: And other laws of cyberspace.* New York: Basic Books.

Lindlof, T. R., & Shatzer, M. J. (1998). Media ethnography in virtual space: Strategies, limits, and possibilities. *Journal of Broadcasting & Electronic Media, 42*(2), 170-189.

McPhee, W. N. (1963). *Formal theories of mass behavior.* New York : The Free Press.

Morris, M., & Ogan, C. (1996). The Internet as mass medium. *Journal of Communication, 46*(1), 39-50.

Netanel, N. (2000). Market hierarchy and copyright in our system of free expression. *Vanderbilt Law Review, 53*(6), 1879-1932.

Neuman, W. R. (1991). *The future of the mass audience.* New York: Cambridge University Press.

Nielsen//NetRatings. (2001, September 10). *netReporter.* Author.

Owen, B. M., & Wildman, S. S. (1992). *Video economics.* Cambridge, MA: Harvard University Press.

Persky, J. (1992). Retrospectives: Pareto's Law. *The Journal of Economic Perspectives, 6*(2), 181-192.

Porter, T. M. (1986). *The rise of statistical thinking: 1820-1900.* Princeton, NJ: Princeton University Press.

Rao, V. R. (1975). Taxonomy of television programs based on viewing behavior. *Journal of Marketing Research, 12,* 335-358.

Roscoe, T. (1999). The construction of the World Wide Web audience. *Media, Culture & Society, 21*(5), 673-684.

Rosengren, K. E., Wenner, L. A., & Palmgreen, P. (Eds.). (1985). *Media gratifications research: Current perspectives.* Beverly Hills, CA: Sage.

Rust, T. R. (1986). *Advertising media models: A practical guide.* Lexington, MA: Lexington Books.

Rust, T. R. , Kamakura, W. A., & Alpert, M. I. (1992). Viewer preference segmentation and viewing choice models of network television. *Journal of Advertising, 21*(1), 1-18.

Vogel, H. L. (1998). *Entertainment industry economics: A guide for financial analysis* (4th ed.). New York: Cambridge University Press.

Webster, J. G. (1985). Program audience duplication: A study of television inheritance effects. *Journal of Broadcasting & Electronic Media, 29,* 121-133.

Webster, J. G., & Phalen, P. F. (1997). *The mass audience: Rediscovering the dominant model.* Mahwah, NJ: Erlbaum.

Webster, J. G., Phalen, P. F., & Lichty, L. W. (2000). *Ratings analysis: The theory and practice of audience research.* Mahwah, NJ: Erlbaum.

Webster, J. G. , & Wang, T. (1992). Structural determinants of exposure to television: The case of repeat viewing. *Journal of Broadcasting & Electronic Media, 36*(4), 125-136.

Williams, R. (1958/1983). *Culture and society: 1780-1950.* New York: Columbia University Press.

Zillmann, D. (2000). Mood management in the context of selective exposure theory. In M. Roloff (Ed.), *Communication yearbook 23* (pp. 103-123). Newbury Park: Sage.