15

Content Diversity Online: Myth or Reality?

Eszter Hargittai
Northwestern University

THE IMPORTANCE OF STUDYING "EXPOSURE DIVERSITY"

A vast amount of material is available to users of the World Wide Web. Search engines index the contents of billions of pages—by last reports one of the most widely used search engines, Google, claims to index over 8 billion—yet they admittedly only cover a fraction of all available content (Sullivan, 2005). Unlike with other media, once content on the Web is publicly available, anyone with knowledge of its Web address can access it. This has led to much enthusiasm regarding the medium, including justification for media deregulation by former Federal Communications Commission Chair Michael Powell (Manjoo, 2003). But does the Web really offer a panacea to concerns about diversity in the media landscape? In addition to looking at what is theoretically available to users, I argue that we must also consider what is realistically within their reach (Hargittai, 2000). To understand realistic accessibility as opposed to mere availability, we must consider what factors influence the types of material that people are most likely to retrieve online. In this chapter, I argue that the way content is organized on Web sites and users' ability to navigate the Web both influence what material is most easily accessible.

Napoli (1999) outlines a helpful framework for approaching questions about media diversity by identifying three different types: (a) source diversity, (b) content diversity, and (c) exposure diversity. Webster (this volume) offers a

helpful summary of this approach and points out the dearth of research on "exposure diversity." His chapter explores the viewing habits of TV audiences in a multichannel environment. In a similar vein, this chapter looks at the actual online actions of Internet users to shed light on whether the promises of the Internet for content diversity are realized when it comes to average users' utilization of the medium.

To explore this question I look at the options for, and processes of, finding information about local cultural events online. The availability of local content is of particular interest for discussions about media diversity as local content can easily be the first to disappear when national media buy up local media outlets. It is cheaper to syndicate a centrally-created program than to run local versions across numerous outlets. However, this may be to the detriment of local information. Can the Web act as an alternative source of information about local events?

I use several methodologies to investigate what options are most realistically accessible to Internet users when it comes to locating information about local cultural events. In the next section, I describe the methods used in the study on which I base the findings in this chapter. Then I look at how content about local cultural events is presented on the most popular Web sites, portals, or "point-of-entry" sites. Next, I discuss findings from the study in which I conducted in-person observations of 100 randomly sampled adult Internet users' online experiences. Results suggest that the factors traditionally responsible for how people are able to access information about movies and other local cultural events are very much mirrored online.

STUDYING PEOPLE'S ONLINE INFORMATION-SEEKING BEHAVIOR

The findings in this chapter are based on an in-depth study of 100 adult Internet users' online information-seeking behavior. In this section, I briefly describe the methods of data collection and analysis. For more details about the methodology, including sampling procedure, recruitment, the interview protocol, particulars about the in-person observations, and description of the coding and classification scheme, please see Hargittai (2002, 2004).

To gain a thorough understanding of people's online information-seeking behavior, I collected data through in-person observations and interviews with randomly selected adult Internet users between the summers of 2001 and 2002 in a New Jersey county. The response rate for participation was 54%, considerably high given that respondents were asked to come to a university location, traveled up to a half hour each way to the study location, and spent on average an hour and a half with the researcher. Study participants reported information about their Web and other media uses on a questionnaire and also supplied basic demographic information.

The main component of the study session involved sitting at a computer and looking for various types of material online. Respondents chose the computer platform and browsing software of their preference. They were then asked to find information ranging from local cultural events to tax forms, information about political candidates, health material, and other content. The questions were presented in sequence one at a time and participants were given unlimited amounts of time to work on each task. The participants were not offered any advice on how to look for content; they had to rely on their existing know-how and experiences. They started out with a blank homepage and it was up to them to find their way to a search engine or any other Web site on which they proceeded to find the requested material. For example, a user may have gone to a search engine and typed in a query to find information. Alternatively, others went directly to a Web site about which they had previous knowledge. A third popular option was to click on links on portal sites and follow the Web site recommendations on the proceeding pages.

Respondents represented a diverse group of Internet users ranging from salespeople to administrative assistants, students to teachers, military personnel, architects, and retired people. They ranged in age from 18 to 81; half were women (51%), and the median level of education was a college degree. Table 15.1 presents specifics about the demographic background of sample respondents. Regarding their Internet experiences, most participants can be considered veterans as the median number of years people had been using the medium was 6 years. Nonetheless, the sample also includes some novice users, with 13% having gone online 2 years within the time the study took place. Participants differed considerably in the amount of time they spent surfing the Web weekly. Although some respondents only went online for a few minutes daily, others were constantly connected, with up to 30 hr of Web use each week. Such diversity among study participants allows us to gain an idea of average users' online experiences instead of limiting findings to a very particular online population. Moreover, we are able to explore the relationship between users' demographic characteristics, their experiences, and their online abilities.

To understand what types of information sources adult Internet users are most likely to consult on the Web, I analyzed participants' online actions. I developed a coding scheme for classifying each move a user can make to go from one Web page to the next. Hargittai (2004) describes this method in detail. The resulting summary documents of people's online actions include every move

TABLE 15.1

Descriptive Statistics about Sample Participants

	Mean	*SD*	*Median*	*Minimum*	*Maximum*
Age	42.96	15.86	42	18	81
Education	16.21	2.72	College	Less than high school	PhD
Family income	$98,394	$57,452	$84,500	$18,250	> $250,000
Number of years since first use of the Internet	6.28	3.38	6	0	16
Number of hours browsing the Web weekly	8.62	9.39	7	8 min	70 hr

participants made with information about the Web sites they visited, how they got there, and how much time they spent on each page. These data allow us to aggregate information about the types of Web sites people visit during their online sessions and how they seek various types of material on the Web. I draw on this nuanced coding of the data to explore later how users look for information about local cultural events online.

In addition to the data described earlier—demographics, user experience, and online actions—I also aggregated information about how much time people spent looking for content and whether they were able to complete tasks successfully. I coded the sessions for completion rate of each task and amount of time spent on each query measured in seconds. These measures reflect people's online abilities and comprise the dependent variable in the analyses shown later. Before presenting the results of the statistical analyses, in the next section I discuss how information about cultural events is organized online.

THE AVAILABILITY OF CULTURAL EVENT INFORMATION ONLINE

Information about movies is quite easily accessible on the Web. Several portal sites have movie links directly on their homepage.[1] In contrast, information about other types of entertainment is much less prominent on the welcome pages of such aggregator sites.

Consider, for example, the layout of categories on Yahoo's homepage as shown in Fig. 15.1. Although there is a "Movies" link right on the homepage, there is no equivalent direct link for those seeking information about theaters or orchestra performances.[2] If a user is looking for movie information, he or

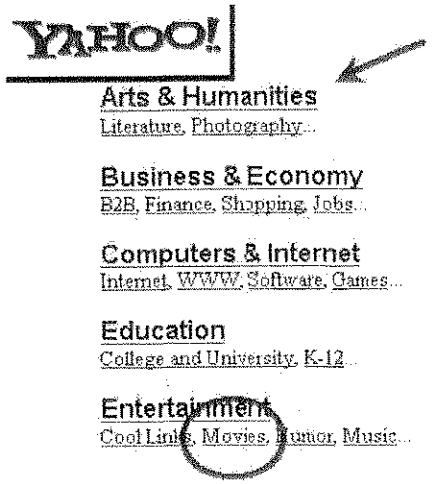

Figure 15.1. Yahoo's directory categories containing movie and theater information, 2001 (time of study). (Arrow and circle added to point out relevant sections of the image.)

she is given the link quite prominently on the homepage. This is the case on other portals used by study respondents as well (e.g., att.net or excite.com). In contrast, those looking for theater information on Yahoo's homepage would have to know that it is located under the directory heading "Arts & Humanities." Moreover, it is not enough to click on that link to access the "Theater category." It takes one more click on the link "Performing Arts" to finally see a link labeled "Theatre."

On one of New Jersey's regional portals, www.nj.com, the "Movies" and "Theater" link were featured at the same level at the time of the research project. This site was popular with several participants of the study. Nine of them used it for finding various types of content. On www.nj.com, both "Movies" and "Theater" links were prominently visible when one clicked on the "Entertainment" link on the homepage (see Fig. 15.2 for details). However, once a user clicked on the "Movies" link she was taken directly to a search engine for movie listings (see Fig. 15.3 for a copy of what the user saw on the resulting

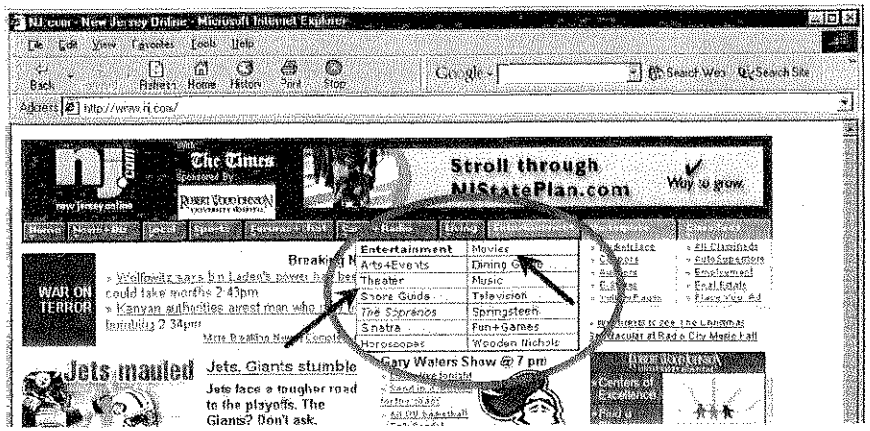

Figure 15.2. The entertainment category links on local Web portal nj.com.

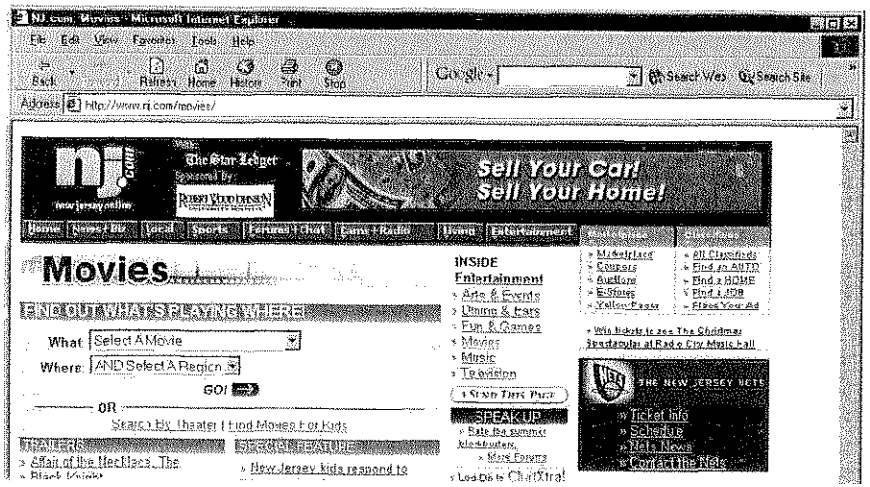

Figure 15.3. Nj.com's Movie search page (2001).

Web page). In contrast, if she clicked on the "Theater" link, she had to guess the next step, because there was no obvious schedule information on the resulting page. If she guessed "right" (picked the second link shown on Fig. 15.4) then she was finally on a page where she could look for theater performance schedules (see Fig.15.5 for the resulting page).

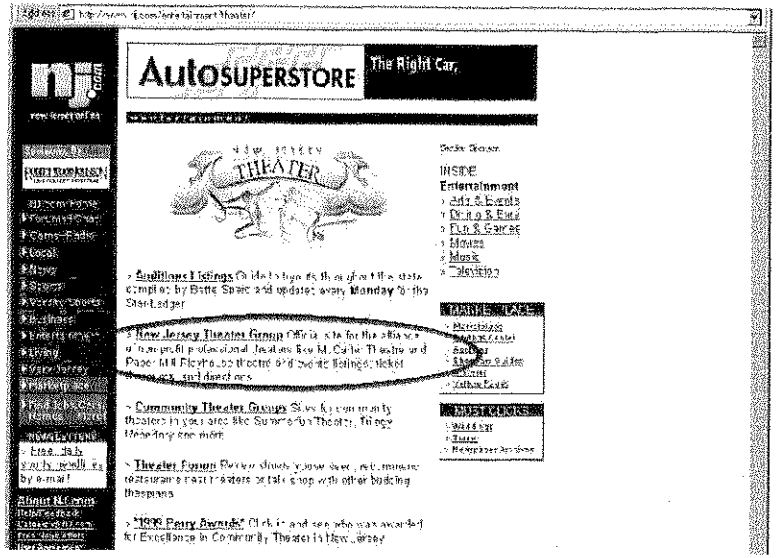

Figure 15.4. The page that comes up after a user clicks on Theater on nj.com's homepage (2001). Circle added for emphasis.

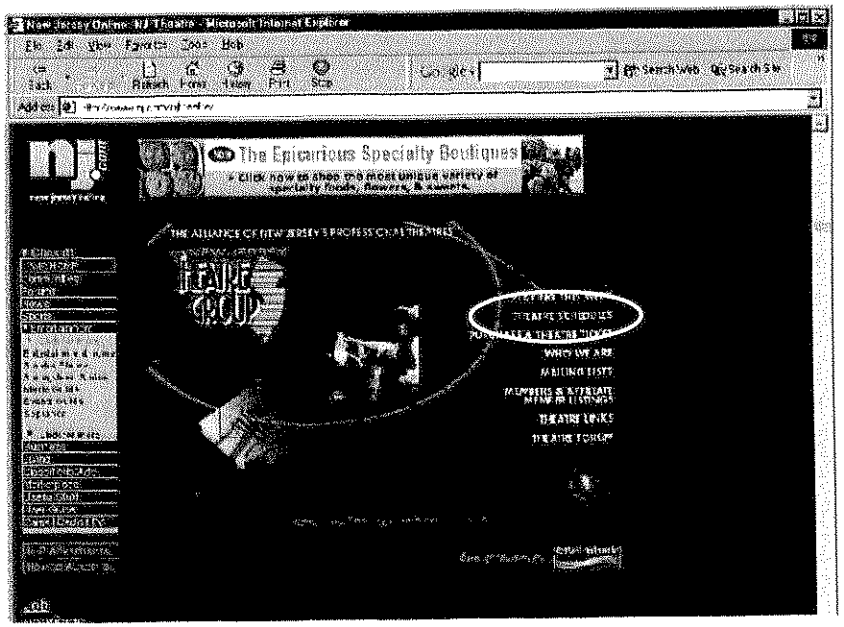

Figure 15.5. Nj.com's Theater page (2001) white circle signaling the "Theatre Schedules" link. Circle added for emphasis.

It is curious that www.nj.com had decided to make it so difficult to get to this page when, in fact, the directory structure of their site as per the page's URL suggests that it was quite high in the directory hierarchy (http://www.nj.com/njtheatre). That is, the fact that "njtheatre" is immediately after the domain name "www.nj.com" signals that the theater directory is a top directory on the www.nj.com site similarly to the movie directory. Again, these images are copies of what users saw in the study in 2001.

A look atwww. nj.com in 2005 shows similar organization of content although a link to theater schedule information is no longer available anywhere on the Web site, making that material even harder to access. The heading "Movies" and "Theater" are still accessible at the same level. However, it is no longer possible to find a page—no matter how far down one tries to drill in the site—that gives the user the option of searching for theater show times in the area. The one seemingly relevant link is specified as an advertisement and links to a page with show times for Broadway musicals. A user may be led to believe that there are no theater performances available in Central New Jersey. This is an incorrect assumption because such award-winning theaters as the McCarter Theatre in Princeton put on performances regularly. Moreover,

McCarter Theatre and other local arts establishments have Web sites of their own publicizing their events. However, without a central searchable repository, they are harder to reach than sites of large venues whose performances are aggregated in large databases.

Another hurdle to finding certain types of content online concerns one's geographical location. Some Web services offer information about events by localities but only aggregate content for metropolitan areas. For example, America Online (AOL) offers localized listings but only for certain metro areas. On the main welcome screen of AOL's interface under the heading "My Places," the user has the option of clicking on the link "Local News" (see Fig. 15.6). The destination page (replicated in Fig. 15.7) shows the list of possible Local News locations. If a user is looking for information about Central New Jersey, then he or she must choose from among New York, North Jersey, or Philadelphia (indicated on the figure with circles). There is, however, no local events page for Central New Jersey.

Although the organization of movie content and information about other cultural events may simply mirror the organization of these cultural industries in the offline world, it is important to recognize that content organization and presentation on the Web for the most part mirror existing industrial organization instead of offering alternative avenues for reaching audiences and allowing them to find detailed local content. Theoretically, there is no reason why a large online database about all national theatrical or symphony performances cannot exist and offer a one-stop access to such information. However, there is no such service and finding information about such cultural events is just as disparate in the online landscape as it is through more traditional methods.

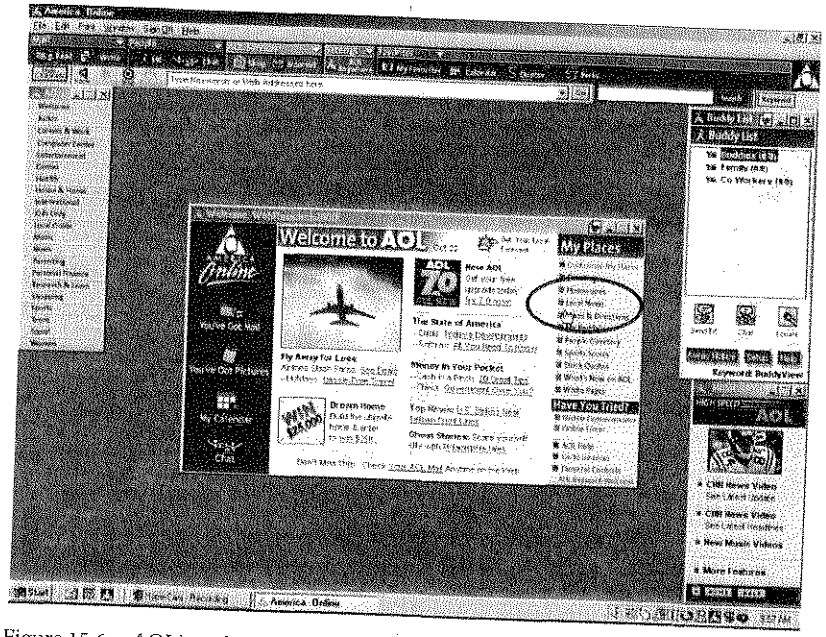

Figure 15.6. AOL's welcome screen with the "Local News" link designated by the circle (added for emphasis).

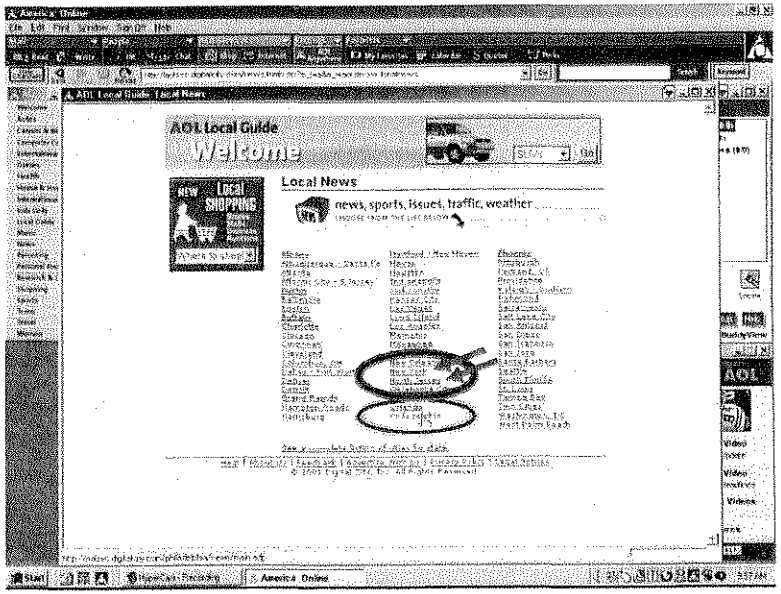

Figure 15.7. AOL Local News options (circles and arrows added for emphasis).

Even if a user knows what steps to take on the portal sites described earlier, he or she is not presented with a national database of theater information in the end. Unlike movie schedules that are aggregated nationally, other types of performances are harder to find online. Users looking for jazz or orchestra performances are also left without a central database to offer information. In the next section, I present the results of an empirical study testing whether the aforementioned differences in how cultural event information is organized online influences people's ability to find local event information.

ACCESSING CULTURAL EVENT INFORMATION ONLINE

The aforementioned description of how cultural information is organized online suggests that movie information may be easier to access than other perfor-

mances due to its more centralized system. Thanks to the research project described earlier regarding users' online information-seeking behavior, it is possible to test this hypothesis empirically. In one of the tasks, respondents were asked to look for information about local cultural events, in particular, the schedule for a movie or theater performance. It was up to participants to decide whether they would look for a movie listing or another type of cultural event such as theater and orchestra performances, symphony or jazz concerts. These specifics were not imposed on participants to explore a task reflective of their own interests.

In the aggregate, 85% of the 100 respondents in the study were able to complete this task successfully. On average, they spent a little more than 2 min on the task. Those who were successful averaged less than 2 min, with the unsuccessful searchers spending over 4 min on this task on average. According to the survey that participants filled out about their earlier Web-use experiences, 61% had looked at a movie or TV Web site in the previous 30 days and 64% had looked at a Web site for music or concert information, so the task was not completely foreign to the majority of users.

Were participants in the sample who looked for movie listings more successful in their quest to find relevant information than those who looked for other types of local cultural events? Results indicate that 9% of those who looked for movie information were unable to find it whereas 33% of those who looked for other types of material were unsuccessful. It seems that indeed, movie information is easier to find online.

More refined analyses are necessary to draw a link between content organization and accessibility of content to users. It may be that certain types of people are more likely to look for movies than others, and those may be the people who are better at searching for online material in the first place. It is indeed the case that older people were less likely to search for movie listings. We know from research exploring age differences among Internet users that older adults spend less time online (Loges & Jung, 2001) and are less likely to find various types of material (Hargittai, 2003). Other demographic characteristics and previous experience with this particular task may also influence people's success with this task, so it is important to account for them when analyzing predictors of task completion.

To test whether it is really the case that movies are easier to find online than other types of local entertainment, I turn to statistical analyses of the data I collected in my study about how users find information regarding local cultural events. I ran a discrete-time logit on hazard for completion of the task (Allison, 1985). I created a person-record file using 10-sec intervals. This new data set allows us to take into account how long people took to complete tasks in addition to information about whether they were successful in finding the requested information. The following is the model specification:

$$\text{Log } (p/1-p) = \beta_0 + \beta_1 \log(t) + \Sigma \beta_j x_j + \varepsilon$$

I use an approximation of Weibull hazard [logged time indicated as log(t) in the model] because theoretically it makes the most sense to assume that there is a log-linear increase in the hazard across time for completing a task (whether that be successful completion or giving up on a task). The model predicts the likelihood of having completed the task in any one time interval.

The model includes information about user demographics, presence of children in the household, quality of equipment, social support networks, autonomy of use, and experience using the Web. There is also a variable controlling for recent experience with this particular task ("looked at movie site"). Additionally, I added a variable signaling whether the respondent had looked for movie time listings as compared to details about other local cultural events. There are two additional dummy variables. They indicate whether the respondent's first move when working on this task was to go directly to a movie site or a theater site. Also, to account for any system-wide changes in the Web during the course of the study, I include a control variable for the amount of time—measured in days—that had elapsed since the first respondent's study session.

The findings presented in Table 15.2 suggest that, controlling for individual characteristics and the conditions of one's use and experience with the task, those who looked for movie listings in response to this task were statistically significantly more likely to complete the task successfully in any one time interval than those who looked for other types of local cultural events. In short, users do a better job in finding movie schedules online than other types of cultural event information. Moreover, those who typed in the URL of a movie site also had a higher rate of completion. In this analysis, I also controlled for having typed in the URL of a theater. This variable, however, does not show any relationship with ability to complete this task. The reason for this may be that theater sites are not always as user-friendly and scheduling information is not as readily available, as on the aggregated movie sites.

Is it generally the case that the traditional outlets are the first point-of-entry for users seeking content online? More than with most other tasks administered in the study, many respondents' first reaction was to look for this type of information in traditional media sources such as the online version of their local newspaper. Among those who looked for movie information in response to seeking local cultural events, 16% of respondents used moviefone.com, 10

TABLE 15.2

Discrete-Time Logit Predicting Hazard for Completion of Task for Finding Information About Local Cultural Events (Standard Errors in Parentheses)

Female	**−0.809****	(0.254)
Age	**−0.035***	(0.014)
Education	0.079	(0.047)
Black	−0.714	(0.429)
Family income	0.115	(0.213)
Child in household	−0.330	(0.314)
High speed	−0.029	(0.332)
Social support	0.042	(0.023)
Net advice	**−0.174***	(0.068)
Free Internet use at work	0.473	(0.266)
No. of Access locations	**1.125****	(0.383)
Time on World Wide Web per week	0.088	(0.201)
Use years	0.203	(0.337)
Looked at movie site	0.242	(0.360)
Movies for task	**0.924***	(0.433)
Direct movie site URL	**0.744***	(0.301)
Direct theater site URL	0.942	(0.750)
Time into study	0.004*	(0.002)
Weibull hazard	0.834***	(0.184)
Intercept	−9.390***	(2.840)
N	1,405	

***$p < 0.001$; **$p < 0.01$; *$p < 0.05$, two tailed.

turned to AMC Theatre's Web site, and seven went to Hollywood.com. That means that over a third of all respondents or 43% of those looking for movies went to a large commercial aggregator site. A surprising number of people looked for such information using local sources. Nine percent of users went to the New Jersey portal www.nj.com, five tried to access the information through the Princeton Packet's online presence, and some tried using a local university's homepage.

CONCLUSION

Although the Internet is undoubtedly a medium like no other in terms of the amount of diverse content it makes available to users, sources closely tied to traditional media and organization of content remain the most easily accessible even in this new environment. In particular, local content that is hard to aggregate into national databases seems to be considerably less within the reach of users than content that is packaged for widespread consumption. The observations presented based on average Internet users' online experiences suggest that mere presence of content diversity online does not guarantee its ease of accessibility.

One possible explanation for how users access content online given the vast amounts of resources is to assume that they gravitate toward their preferences. Given the myriad of information sources online, we may expect that users choose whatever they deem most relevant. If this is the case, then users are simply exercising their choice by going to sites that mirror traditional media. However, it would be wrong to assume that the content users access is necessarily a reflection of their preferences. Several factors such as online content organization and user skill are at work when users browse material on the Web and these influence what content people are more or less likely to access.

The most popular portal sites represent enormous media companies (Manjoo, 2003). Although it is not necessarily the case that they are all owned or affiliated with big traditional media companies, they are certainly parts of significant ventures. The amount of advertising and resources that go into these services allows them to far surpass what a small unaffiliated Web site can achieve in popularity and exposure. Another factor that influences what types of content users reach concerns their online abilities. Users differ considerably in how well they are able to navigate online materials. Those who are more skilled will be more likely to find the types of content of direct relevance to their interests. However, others will be more dependent on the information presented on easily accessible sites such as big portals.

Overall, it would be wrong to assume that the mere presence of diverse material on the Web will result in users accessing a smorgasbord of content. As presented in this chapter, an interplay of many factors determines what information is most realistically within the reach of users. It is important to draw a distinction between available and accessible content online. When we assess the state of online content diversity, we must rely on data about users' actual behavior in addition to considering what content exists on the Web. As shown in this chapter, Internet use can reflect offline content utilization regardless of what is theoretically available to users.

NOTES

[1]By "portal sites" I refer to point-of-entry sites such as Yahoo! or MSN that often are set as default browser pages by Internet service providers and constitute the first page users see when they go online.

[2]The illustrations shown in this chapter are from the time of the study, 2001 to 2002, to reflect what study respondents saw when they participated in the study. Yahoo's

Web site directory today, in 2005, also prominently displays a "Movies" category right on the homepage while continuing to have no direct link called "Theater."

REFERENCES

Allison, P. (1985). *Event history analysis.* Thousand Oaks, CA: Sage.

Hargittai, E. (2000). Open portals or closed gates? Channeling content on the World Wide Web. *Poetics, 27, 233–253.*

Hargittai, E. (2002). Beyond logs and surveys: In-depth measures of people's Web use skills. *Journal of the American Society for Information Science and Technology Perspectives, 53,* 1239–1244.

Hargittai, E. (2003). *How wide a web? Inequalities in accessing information online.* Unpublished doctoral dissertation, Princeton University, Princeton, NJ.

Hargittai, E. (2004). Classifying and coding online actions. *Social Science Computer Review, 22,* 210–227.

Loges, W. E., & Jung, J.-Y. (2001). Exploring the digital divide: Internet connectedness and age. *Communication Research, 28, 536–562.*

Manjoo, F. (2003). Can the web beat big media? *Salon.* Retrieved November 20, 2004, from http://www.salon.com/story/tech/feature/2003/054/21/web-vs-big-media/ index-np.html

Napoli, P. (1999). Deconstructing the diversity principle. *Journal of Communication, 49*(4), 7–34.

Sullivan, D. (2005, January 28). Search engine sizes. *Search Engine Watch.* Retrieved October 13, 2005, from http://searchenginewatch.com/reports/article.php/2156481

16

Proactive Media Policy in an Age of Content Abundance

Ellen P. Goodman

Rutgers–Camden School of Law

Imagine a mediascape in which public appetites for video programming were completely satisfied. There was enough video on demand, broadband videocasting, peer-to-peer distributed video, as well as scheduled cable, satellite, and broadcast programming to satisfy the tastes of most every audience member, whether those tastes ran to conservative commentary, Spanish-language dramas, reality programming, or local news. Would this be a mediascape in which the aims of media policy were satisfied?

The central question in media policy debates today is how digital networks and consumer technologies affect the pursuit of traditional policy goals like localism and diversity. With such an abundance of media content available over cable, satellite, and the Internet, and with increased consumer control over access to such content, is government intervention in media markets still necessary?

For the most part, policymakers and advocates begin to address this question from a seductively uncontroversial premise: that the goal of media policy is to protect public access to a diversity of speech, including "local" and noncommercial content (Owen, 2003). It is from this premise that arguments typically flow about media concentration, public interest obligations, and the need for public broadcasting. So framed, these arguments turn on disputes about how well media markets function—essentially, about the responsiveness of these markets to smaller taste and geographic communities.

Lawrence Erlbaum Associates, Inc., Publishers
10 Industrial Avenue
Mahwah, New Jersey 07430
www.erlbaum.com

Cover design by Kathryn Houghtaling Lacey

Library of Congress Cataloging-in-Publication Data

Media diversity and localism : meaning and metrics / edited by Philip M. Napoli.
p. cm. — (New directions in communication disorders research, integrative approaches)

Includes bibliographical references and index.
ISBN 0-8058-5548-3 (cloth : alk. paper)
ISBN 1-4106-1440-9 (e book)
1. Mass media—Ownership. 2. Local mass media. 3. Mass media and minorities. 4. Mass media policy. I. Napoli, Philip M. II. Series.
P96.E25M39 2006
302.23—dc22 2006003198
CIP

Books published by Lawrence Erlbaum Associates are printed on acid-free paper, and their bindings are chosen for strength and durability.

Printed in the United States of America
10 9 8 7 6 5 4 3 2 1

Contents