

Running head: ASSESSING VIRTUAL DIVERSITY

Assessing Virtual Diversity

Project Proposal

Bradley Osborn

Mass Communication Research Methods

Instructor: Dr. Rick Fischer

The University of Memphis
Memphis, Tennessee

Fall 1999

Style: Publication Manual of the American Psychological Association. (4th ed.). (1994).
Washington, DC: American Psychological Association.

Strunk, W. (rev. by White, E.B.). (1999). The Elements of Style, (4th ed.). Boston:
Allyn & Bacon.

CONTENTS

Introduction	4
Problem Statement	4
Purpose of the Study	5
Literature Review	5
Current Resources and Research	5
Market-Driven Research	6
Government and Academic Resources	7
Advocates	8
Method Review	9
Research Questions	11
Methodology	11
Preliminary Sample Data	13
Statistics	13
References	14
Appendix 1 – Survey Instrument	16
Appendix 2 – Cost	21
Appendix 3 – Timeline	22
Appendix 4 – Email Invitation	23

Introduction

Problem statement

As the global population moves past six billion, estimates of its composition vary, but one thing is certain: It will not remain static. In 2050, the United States will rank a distant third among nations with the greatest populations. Projections have India leapfrogging China to occupy the top spot. That same year, America will have become one quarter Hispanic. In addition to geography and ethnicity, the elder population will soar. Those with disabilities and sexual minorities will have voice. Rural, religious, and youth components along with the economically disenfranchised, will also be present.

(The World At Six Billion, 1999)

One conceivably unifying and universally empowering force will be the medium that has come to be known as the Internet, or the World Wide Web. Once the bastion of bureaucrats, scientists, the military, and academia, the Internet is now a potential resource for all. Ideally, these divisions among peoples should dissolve in the virtual world, but they remain topics of interest to marketers, demographers, researchers, and advocates of civil, equal opportunity, and economic rights.

However, the Internet affords an unprecedented degree of anonymity, as it should. Until these divisions truly dissolve in the real world, they possess unique attributes sought by others to whatever end. To further these ends, especially legitimate research, practitioners must find means to accurately measure this population, its makeup, and its degree of diversity. Although Internet surveys, polls, market analyses, and customer profiles are ubiquitous, the opacity of the virtual window remains.

Purpose of the Study

This study will make an elementary assessment of Internet demography based upon active subscribers of one Internet service provider (ISP) or Internet access provider (IAP). Subsequent studies of other providers may amplify and/or augment the data and results produced by this project.

Literature Review

Current Resources and Research

Undoubtedly there exists a wealth of proprietary data that relate to web diversity, gathered for commercial purposes, using retailer surveys, customer profiles, and other market tools. Government and academia are less-restricted sources for such information. A third entity class, possessing far fewer resources for data acquisition, but with certain vested interest in doing so, is one made of groups advocating civil liberties, human, equal, and economic rights.

The archetypal Internet merchant – if there is such a thing -- appears to be ahead of the fray in gathering virtual demographics, whether obtained directly from web patrons, third party agents, or procured on a subscription basis. Readily accessible customer bases and corporate underwriting fuel their efforts. However, the fruits of these corporate endeavors are generally private, and fall short when held to scientific scrutiny, their goal being enhancement of a bottom line, not edification or altruism. ISP/IAPs are also in possession of the much-sought-after subscriber profile.

Market-Driven Research

Early 1999 combined Business Week and Harris polls show that 46% of those who use the Internet are women, up 6% from 1996. According to Business Week (1999, March 22), Forrester Research's year-end US forecast has 32% of black, 43% of Hispanic, and 67% of Asian-American households online, compared with 39% of white households. The percentage of wired Europeans should rise from 4% to 13% by 2001. The number of Chinese on the net will more than quadruple, to 9.4 million, by 2002. Usage rates by those aged 55 to 64 will grow to 40% by 2003 (Crockett, 1999).

English speakers comprise 57.4% of Internet users, while those speaking other European tongues make up 26.4% (Global Reach, 1999).

Research firm eMarketer's new eUser & Usage Report, predicts the total Internet population in the United States, including children, will reach 80.8 million by year-end 1999. The Usage Report, a product of secondary research on US samples, contains extensive Internet market data characterizing the American net user. Their typical user profile: a 38-year-old educated, affluent male. (eMarketer, 1999)

Key findings from the eUser & Usage Report:

- The median income for online households is 57% higher than that of the average American household -- \$58,000 vs. \$37,005, respectively.
- Some 27 million women now online account for 46% of all U.S. net users.
- Teens average 8.5 hours online per week -- 27% more than average net users.
- 87% of college students are currently online, representing, by far, the most active single group on the net.
- In 1999, seniors will account for 19%, or \$3.5 billion, of total consumer online spending in the U.S. With a compound growth rate of 68%, this figure will reach \$16.7 billion by 2002.
- Hispanics represent the largest minority group online with 6.9 million net users, followed by 6.6 million blacks, and 4 million Asians.
- The number of gays and lesbians online worldwide will rise from 9.2 million in 1999 to 17 million in 2005.

Other market-based sources include the CommerceNet/Nielsen Media Research June '98 Internet Demographic Survey, and the Willard & Shullman Group Ltd. (WSG, <http://www.wsgresearch.com/>). According to Business Wire, November 10, 1999, mass market companies are starting to target African-Americans and gays, but not the Asian, Christian, Latino, and female online communities (Business Wire, 1999). All of these findings and services are heavily slanted toward the American user.

Government and Academic Resources

The United States government is a grand repository of *démographie sociale*, especially the constitutionally mandated Census Bureau, a source for current US demography of any sort. However, much of the government data flow from the Department of Commerce, and as such are tied to our pocketbooks. In addition, exclusive use of these palettes would be to paint with a particularly Americentric brush.

University scholars have provided little data on web diversity. Nonetheless, collegiate Internet user lists, although limited in scope, are not-so-shallow wells from which to draw.

As per the Trendline Study On Electronic Access By Households: 1984-1998 from the Bureau of the Census, those Americans enjoying the greatest connectivity today are typically those within high-income households. Holding income constant, other highly-connected groups include whites or Asians, middle-aged, highly-educated, employed, and/or married couples with children, most often found in urban areas and the West. (Bureau of the Census, 1999)

Conversely, the least connected generally are low-income, black, Hispanic, or Native American, senior in age, not employed, single-parent (especially female-headed) households, those with little education, and those residing in central cities or especially rural areas. (Bureau of the Census, 1999)

Advocates

Although quick to exploit findings of inequity in the use of the world wide web, champions of egalitarianism are not yet a major source for the assessment of virtual diversity.

- A third of African Americans online only have access outside the home, be it at work, the library, at school, or at a community center (Wagner, 1999).
- Along with AT&T, the National Association for the Advancement of Colored People (NAACP) has launched a \$300,000 program to get a larger diversity of people surfing the Internet (Computing, 1999).
- Among those earning \$15,000 to \$35,000 a year, more than 32% of white families owned computers, but only 19% of black and Hispanic families had computers at home. That gap widened from eight percentage points five years ago even as the price of entry-level personal computers plunged (Sanger, 1999).

- Within a week of the first NATO bombs dropping on Serbia, the conflict in Kosovo had been dubbed the first Internet war. While that overstates the case, it is not far from the truth (Asica, 1999).
- The Internet access gap between the U.S. and the rest of the world remains (Stewart, 1999).
- A growing number of Americans are reporting some level of disability, according to a Census Bureau report. And the figure is sure to get larger yet as baby boomers age (Stoneman, 1998).
- The number of Americans aged 65 and older increased from 28 million in 1985 to 34 million this year, and is projected to reach 39 million by 2010 (Edmondson, 1997).

It is the purpose of this project to make an elementary assessment of this e-diversity, with as much detachment and objectivity as is practicable, based not on market, academic/government, or human-interest data. Instead, the information will come from a random sample taken from the population of active users of one or more global providers of Internet access. Active users are those user currently logged onto the Internet.

Method Review

Internet service providers and Internet access providers usually allow subscribers to complete a customer profile. This profile can hold any number of demographically useful pieces of data. Although, itself, a tempting source, this resource is biased by variability in completeness, multiple screen identities, and user embellishment. Instead, for this project, these user profiles shall be queried for a random word or phrase, at several time points, over several days, without regard to any predetermined demographic parameter(s).

Inherent bias exists because the query and project are to be conducted in English. There are also no ISP/IAPs with truly global subscribership. Therefore, several different provider client bases should serve as populations for subsequent

studies, and ISP/IAPs providing, both paid and free access, will ultimately be surveyed, including those who gain access through employers, public facilities, and educational institutions.

The sample will be an email cluster sample, and will attempt to be representative of the global Internet population. An initial and reminder email survey will comprise the first wave.

The cluster will be defined as those persons signed on via a large worldwide ISP/IAP, e.g., AOL, at certain predetermined times of the day, in a twenty-four hour period, over several days. Should the list become unmanageably large, a further random reduction device will be employed. Soon (days) after the cluster is defined, the surveys will be emailed to potential respondents.

Follow up and verification activities (second and third waves) could include telephoning, mailing (post), or physically calling upon respondents.

Couper, et al (1997) found an email survey obtained an average response rate of 42.6%, and in Parker's 1992 study of AT&T employees, she reports a 63% response rate for email surveys (Schaefer & Dillman, 1998).

Additionally, commercial email survey providers such as SurveyBuilder.com offer web-based services to this end.

Research Questions

Survey questions will cover Internet access, age, citizenship, disability, education, employment, gender, geographic residence, income, language, marital status, political affiliation, race, religion, sexual orientation, and size of community. The questions and answers are generally presented alphabetically, and there are no preselected answers.

Citizenship, language, and race categories are partially based upon United Nations ([The World At Six Billion](#), 1999) and Global Reach (1999) projections and data.

What is the true demography of the Internet as determined by this method? How do these findings compare to market-based and government assessments? How does American Internet dominance skew these findings?

Justification for Internet diversity studies lies in the pursuit of unidentified and untapped niches or demographic elements within the global wired community. These groups are potential targets for advertisers, rights groups, government and academia. Studies of this nature can also aid in the characterization of those without access to the Internet.

A meta-analysis of current research is an optional adjunct to this study.

Methodology

A random word generator program, such as Random Word Generator v.2 from Gammadyne Software, will be employed to obtain a set of 50 unique English words. If the chosen program generates nonsense words, as well as legitimate English words,

then the first 50 unique English words which pass muster against Merriam-Webster's Collegiate Dictionary Tenth Edition will be used as potential query strings.

After obtaining the query strings, a given Internet access or service provider's user directory will be searched for active users whose profiles include these words. This will be done within a 30-minute window at 1:00 AM, 7:00 AM, 1:00 PM, 7:00 PM in the researcher's time zone. This will be repeated each day for seven consecutive days.

Should any single query produce greater than 100 results, or be deemed too broad, the subsample will be discarded, and the next English word will be used to search the user database. This process will continue until ten words have successfully produced a result, or the 50 word query pool is exhausted. Any duplicate users returned at anytime during the seven-day period will be discarded without replacement.

Sample members will be emailed the invitation to participate in the study during the week after the sample acquisition. A follow up invitation will be sent two weeks after that. Respondents will not be allowed to submit more than one survey, and will be given an assurance of anonymity. Respondents will be given the option to provide contact information for subsequent telephone, mail or physical calls. Respondents are under no obligation to provide this information. If they choose to do so, the anonymity assurance is voided, and an assurance of confidentiality is given in its place.

Preliminary Sample Data

A preliminary trial of the sample acquisition portion of the method at one time-point (1/28 of total sample acquisition) yielded the following:

<u>Random English Word</u>	<u>Active AOL Users Returned</u>
rice	87
ego	45
tap	48
may	too broad
life	over 100
phyla	no matches
jam	over 100
sunk	no matches
gal	over 100
treed	no matches
use	over 100
ally	53
among	no matches
ore	16
hake	no matches
orate	1
ruling	over 100
dog	over 100
got	no matches
tin	70
rem	23
arse	16
awe	19
Total	383

Statistics

Data will be analyzed applying all appropriate statistical devices, with the aid of George Relyea, Academic Computer Services, The University of Memphis.

References

- Asica, J.D. (1999, June). Conveying the war in human terms. American Journalism Review, 76.
- Bureau of the Census. (1999). Trendline Study On Electronic Access By Households: 1984-1998 (Appendix to Falling Through The Net: Defining The Digital Divide). Washington, D.C.: US Government Printing Office.
- Business Wire, (1999, November 10). Mass market companies are starting to target African-Americans and gays but not the Asian, Christian, Latino, and female online communities (Computer Economics survey) [Online]. Available: <http://businesswire.com/>
- CommerceNet and Nielsen Media Research. (1998, June). CommerceNet/Nielsen Media Research Internet Demographic Study. [Online]. Available: <http://www.nielsenmedia.com/interactive/commercenet/S98/>
- Couper, M. P., Blair J., & Triplett T. (1997). A comparison of mail and e-mail for a survey of employees in federal statistical agencies. Paper presented at the annual conference of the American Association for Public Opinion Research, Norfolk, VA.
- Crockett, R. (1999, March 22). Special report: e.biz. Business Week (EB46), 3621.
- Edmondson, B. (1997, June). A better chance for healthy old age. Forecast, 17, (6), 1-2.
- eMarketer. (1999, October 19). eUser & Usage Report (Vol. 1 abstract). [Online]. Available: <http://www.emarketer.com/>
- Global Reach. (1999, November 11). Global Internet statistics (by language) [Online]. Available: <http://www.greach.com/globstats/>
- In brief. (1999, July 22). Computing, 46.
- Parker, L. (1992). Collecting data the e-mail way. Training and Development, 52-54.
- Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat. (1999, October 12). The World At Six Billion. (ESA/P/WP.154). New York: United Nations.

Sanger, D.E. (1999, July 9). Big racial disparity persists in Internet use. New York Times, p. A12.

Schaefer D.R., & Dillman, D.A. (1998, Fall). Development of a standard e-mail methodology: results of an experiment. Public Opinion Quarterly, 62,(13), 378-379.

Stewart, T.A. (1999, July 5) A nation of net have-nots? No: Internet ubiquity. (staying smart/managing). Fortune, 140,(1), 184+.

Stoneman, B. (1998, January). Ranks of disabled growing. Forecast, 18,(1), 2-3.

Wagner, V. (1999, November 12). Black Americans surge onto net , fastest growing group to get wired. San Francisco Examiner,. p. A1.

Appendix 1 Survey Instrument

Assessing Virtual Diversity

<http://www.people.memphis.edu/~bosborn/avdsurvey.htm>

Please take a few moments to complete this survey.
Anonymity is assured[†].

1. Do you have current, ongoing browsing/email capabilities via the Internet?

- I can **browse**, and **email**
- I can **browse**, but not email
- I cannot browse, but can **email**
- I cannot browse, or email

2. What is your age?

3. What is your citizenship?

- American
- Bengali
- Brazilian
- Canadian
- Chinese
- European Union
- Indian
- Indonesian
- Nigerian
- Pakistani
- other

4. Are you a person with a disability?

- no
- yes

5. What is the highest level of education that you have completed?

- K-4
- 5-8
- high school
- college
- graduate/professional
- post graduate/specialty

6. What is your employment status?

- full-time (40+ hours per week)
- part-time (1-39 hours per week)
- retired
- student (7+ hours current enrollment)
- unemployed

7. What is your gender?

- female
- male
- transgendered

8. In what geographic region is your primary residence?

- Africa
- Asia
- Australia
- Europe
- North America
- Oceania
- South America

9. Characterize your annual income. ([currency converter](#))*

- high (> 100,000 USD)
- low (<15,000 USD)
- middle (15,000-100,000 USD)

*<http://www.xe.net/ucc/>

10. What is your primary language?

- Chinese
- English
- French
- German
- Italian
- Japanese
- Korean
- Portuguese
- Scandinavian
- Spanish
- other

11. What is your marital status?

- divorced
- married
- partnered
- single
- widowed

12. What is your political affiliation?

- Communist
- Conservative
- Democrat
- Liberal
- Libertarian
- Reform
- Republican
- Socialist
- other
- none

13. What is your race?

- aborigine/native
- Arab
- black
- Chinese
- Indian
- Japanese
- Korean
- Latino
- white

- biracial
- transracial
- other
- unknown

14. What is your religious affiliation?

- Buddhist
- Christian
- Hindu
- Islam
- Judaism
- pagan
- other
- none

15. What is your sexual orientation?

- bisexual
- heterosexual
- homosexual
- prefer not to say

16. In what sized community do you reside?

- large metropolis (>500,000)
- rural (not incorporated)
- small metropolis (20,000-500,000 isolate)
- small town (<20,000)
- suburb (20,000-500,000 adjacent)

17. What is your primary Internet access location?

- friend/relative account
- my account (home)
- public institution
- school
- work
- other

† Providing the following information is VOLUNTARY. Should you decide to provide the researcher(s) with contact information, the assurance of anonymity is voided. However, any contact information that you provide will be held as confidential, and be known only to the researcher(s) participating in this project.

first name

last name

street address

state/province

postal code

country

email

Appendix 2 Cost

Item	Detail	Cost (USD)
Internet service/access provider subscription	AOL, AT&T, CompuServe, EarthLink, MCI WorldCom, MSN, etc.	0.00-26.95
Merriam-Webster's Collegiate Dictionary, 10 th Ed.	Bound and CD-ROM	0.00-39.95
Random Word Generator	Software version 2.0	0.00-15.00
Stationery	printer paper	0.00-2.20
Spreadsheet/statistical software	Excel/SPSS	0.00-249.00
survey and form agents	Formsite.com/ Surveybuilder.com subscription	0.00-9.95
Total		0.00-343.05

Appendix 3 Timeline

Action	Time
Sample acquisition	One week
data acquisition	One month
data analysis	One month
project write-up	Three months
Total	~six months

Appendix 4 Email Invitation

Subject: survey assessing virtual diversity
Date: DD-MM-YYYY
From: bosborn@memphis.edu (Bradley Osborn)
To: randomlychosenuser@ISP.com

You have been chosen to participate in a survey designed to measure the demographic makeup of Internet users.

Your participation is voluntary, and anonymity is assured.

Please take a few moments to follow the link below.

<http://www.people.memphis.edu/~bosborn/avdsurvey.htm>

There are seventeen questions, which can be answered with the click of a mouse and a couple of keystrokes. About nineteen clicks and you are done.

If you want your “group” represented in this survey, just follow the link!

Bradley Osborn
graduate student
Department of Journalism
The University of Memphis
bosborn@memphis.edu