Association for Information Systems
AIS Electronic Library (AISeL)

December 2003

Investigating Online Consumer Behavior Using Thin Slices of Usability of Web Sites

Ananth Chiravuri
University of Wisconsin, Milwaukee

Laura Peracchio
University of Wisconsin, Milwaukee

Follow this and additional works at: http://aisel.aisnet.org/amcis2003

Investigating Online Consumer Behavior Using Thin Slices of Usability of Web Sites

Ananth Chiravuri
University of Wisconsin, Milwaukee
ananthc2@uwm.edu

Laura Peracchio
University of Wisconsin, Milwaukee
lperacch@uwm.edu

Abstract

Studies in psychology have indicated that 'thin slices,' very brief exposure to information, have a great deal of predictive validity in judgments and decision-making. We apply this concept of "thin slices" to study the behavior of consumers when surfing online. Our primary objective is to investigate whether consumers' thin slice judgments of web sites have validity in an online environment. We determine predictive validity of thin slices using measures on the overall usability of a web site and its navigability. We propose to test the hypotheses by conducting an experiment.

Keywords: E-commerce, online consumer behavior, thin slices, Web site design

Introduction

According to the latest U.S. Department of Commerce report, total e-commerce sales for 2002 are estimated to have grown to $45.6 billion, an increase of 27% from 2001 (www.census.gov). Studies conducted on why consumers shop online reveals that a typical online shopper is time starved and therefore uses the web to save time (Bellman et al. 1999). This has generally translated itself into an increase in the time spent online by shoppers. At a broad level, the time spent online by web surfers (including shoppers and non shoppers) has increased because of the surfers' increased familiarity and dependence on the Internet and the growth of high-speed Internet (**http://www.nielsen-netratings.com**). While the total time spent online by web surfers is increasing, there has also been an increase in the number of unique web sites (**http://www.oclc.com**) that has resulted in a matching increase in the number of page views. The net impact of the increase of page views is that the average time spent per page view by surfers in the United States is approximately constant at fifty-five seconds (**http://www.nielsen-netratings.com**).

One of the implications of the data reported above is that consumers may be forming judgments about the web site, the vendor and the products and services offered by the vendor by spending, on an average, less than a minute at each page. These judgments are called "thin slice judgments" or 'thin slices'. Once such judgments are made, it provides the anchor from which subsequent judgments are realized (Ambady et al. 2000b). When applied in the context of an online environment, it is possible that thin slices about overall usability formed from the elements of a given web site could be playing a role in determining the actions of a consumer. For example, online consumers could use their "thin slices" of overall usability of a web site to decide whether or not to shop at that site.

Although there have been numerous research studies on online shopping looking at the different factors affecting it such as gender (Van Slyke et al. 2002) and culture (Chau et al. 2002), there are practically no studies that have looked at the role of thin slices in an online context. This should be an area of further attention because past research indicates that thin slices lead to a great deal of predictive validity in a variety of social situations (Ambady 1999; Babad et al. 1991; Prickett et al. 2000). Examining online shopping from the perspective of thin slices could provide yet another explanation as to why the design of a web site is a crucial determinant of whether visitors are likely to return to the site (Klein 1998). In particular, the objective of this study is to examine whether thin slice judgments of consumers made in the context of online surfing have predictive validity regarding the overall usability of the web site. We measure usability using the "overall usability scale" developed by Agarwal and Venkatesh (2002).

The rest of the paper is structured as follows. We begin by specifying the operational definitions of the variables that will be used in the study. We then describe our hypotheses and the methodology that will be used to test these hypotheses. Finally, we conclude with a general discussion on the contributions of this study.

Operational Definitions

Thin Slices

A thin slice has been defined "as a brief excerpt of expressive behavior sampled from the behavioral stream" (Ambady et al. 2000b, p.203) where a brief excerpt is any excerpt less than five minutes long. For the purposes of this paper, we use this definition to define thin slices as it is appropriate and comprehensive. Thin slices can be sampled from any available channel of communication, including the face, the body, speech, the voice, transcripts or combinations of the above. (Ambady et al. 2000b, p.203). We argue that thin slices can also be sampled from the Internet because it is and will continue to be as important a communication channel as television, sharing most of the latter's characteristics with the added advantage of personalization.

Studies on the process of thin slices indicate that people's perception consists of two stages: the first stage is the automatic evaluative stage that involves minimal cognitive processing and the second stage is the more controlled deliberative stage that is marked by more elaborate cognitive processing and effort (Anderson et al. 1996; Gilbert and Krull 1988; Trope 1986). Present research indicates that the initial evaluative stage is likely to be more prominent than the second stage in the formation of thin slice judgments because of "a combination of the brevity of the stimuli and the nature of the information being conveyed" (Ambady et al. 2000a, p.230).

Thin slices have demonstrated its predictive validity in a number of different contexts. For example, thin slices have been found to reveal teachers expectancies of students (Babad et al. 1991) and also reveal teachers biases. Thin slice judgments revealed that biased teachers addressed their classes in a warmer and less dogmatic fashion verbally but in a less warm and more dogmatic fashion non-verbally than did teachers scoring low on bias (Babad et al. 1989). In addition, thin slices have also been used to provide valid information regarding teacher effectiveness (Stalling and Spencer 1967), predict student achievement (Ambady 1999), examine job performance (Hecht and LaFrance 1995), predict camp counselor effectiveness (Blanck and Rosenthal 1984), predict the outcome of a structured employment interview (Prickett et al. 2000), predict doctor's effectiveness in their referral of alcoholic patients (Milmoe et al. 1967) and to predict patient satisfaction with their doctors (Hall et al. 1981).

Thin slice judgments are equally effective in the case of relationships. Thin slices have been found to differentiate characteristics of dyadic relationships (Ambady et al. 2000a). The findings from their study support the notion that perceptions and impressions of interpersonal relationships are formed accurately almost instantaneously.

Finally, thin slices have been used to reveal individual differences such as personality traits (Borkenau & Liebler 1992; Funder and Sneed 1993; Kenny 1994), gender related features (Bernieri et al. 1992; Cutting and Koslowski 1977; Frable 1987; Lippa 1978) and to detect sexual orientation (Ambady et al. 1999).

These preceding studies from different streams give support to the predictive validity of thin slices judgments in a variety of social situations and form the motivation to explore the issue of thin slice judgments in the context of online surfing.

Usability

Usability and utility are subcategories of the more general term "usefulness" [Grudin 1992]. "Utility is the question of whether the functionality of a system can, in principle, support the needs of users, while usability is the question of how satisfactorily users can make use of that functionality" (Hilbert and Redmiles 2000, p. 388).

Usability has been conceptually defined and operationally measured in multiple ways (Agarwal and Venkatesh 2002). According to Nielsen (1993), usability has multiple components and is traditionally associated with five usability attributes such as learnability, efficiency, memorability, errors, and satisfaction. Another definition advanced by Lecerof and Paterno (1998) includes attributes such as relevance, efficiency, learnability, safety of the system and finally the users' attitude to the system.

For the purposes of this paper, we use the ISO definition as it is appropriate and suitable for our study and define usability as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" (Karat 1997, p.34; Agarwal and Venkatesh 2002, p.170). We measure usability by the three item scale developed by Agarwal and Venkatesh (2002) (see Table 1 for the items).

Table 1. Items Used to Measure Overall Usability (Adapted from Agarwal and Venkatesh 2002)

How do you rate the overall usability of the Web site?

How do you rate the overall design of the Web site?

How do you rate your overall experience at the Web site?

Accuracy

For the purpose of this study, we intend to measure "thin slices" by the accuracy of judgments of web surfers. Using the guidelines from Ambady and Rosenthal's (1992) work on thin slices, we begin by clarifying what we mean by accuracy for this study. Previous research describes accuracy as a correspondence between a judgment and a criterion (Brown 1965). We follow the existing approach in the literature (Ambady and Rosenthal 1992) and use this definition of accuracy for our study too because it implicitly includes the concept of reliability. In this study, we measure accuracy by collecting and correlating criterion ratings and the subjects' ratings of a site.

To begin with, the criterion scores for the web sites are collected by having experts evaluate them specifically on the overall usability as mentioned earlier. Self reported measures are excluded as criterion variables because they may be influenced by factors such as social awareness and self-knowledge of the target individual (Cheek 1982). We then run the experiment and collect ratings of those sites from the subjects. These ratings are then correlated with each other to measure the 'accuracy' of the subjects' ratings. Accuracy, for the purposes of this study, is thus defined as the correlation between the judgments of a subject with the criterion judgments (experts).

Hypotheses and Methodology

Hypotheses

Since the average time per page view is very limited because of the large number of web sites, we hypothesize that consumers operate on thin slices when surfing online. A consumer could form an overall opinion of the overall usability of a web site that would then decide his/her consequent actions. As noted in the earlier section, since thin slices lead to fairly accurate judgments when predicting individual performance, we hypothesize that an online consumer's thin slice judgments of the overall usability of a web site would be as accurate as the longer term judgments of experts. Otherwise stated,

> *H1: An online consumer's thin slices of overall usability of a web site will lead to accurate judgments.*

Previous research has indicated that the accuracy of thin slice judgments does not depend on the length of the behavioral stream. "Observer judgments that were made with slices under 30 seconds in length were as accurate as judgments based on slices nearly five minutes in length" (Ambady et al. 2000b p. 248; Ambady and Rosenthal 1992). Following this, we examine whether this will also hold for consumers surfing online and hypothesize that there should be no significant statistical difference between thin slices of overall usability for different levels of exposure (low, medium and high). In other words, we hypothesize that thin slices of overall usability should be accurate for any level of exposure and state it as:

> *H2: There should be no difference in the accuracy of an online consumer's thin slices of overall usability of a web site for different levels of exposure: low, medium and high.*

Usability research has also indicated the navigability and organization of a web site as important variables (Nielsen 2000; Shneidermann 1998). For the purposes of this study, we use the existing definition of navigability and define it as "the sequencing

of pages, well organized layout and consistency of navigation protocols," (Palmer 2002, p. 155). Following hypothesis 2, we propose that for different levels of exposure, the accuracy of an online consumer's thin slices on the navigability of a site should be the same. This is stated as follows:

> ***H3: There should be no difference in the accuracy of an online consumer's thin slices of navigability of a web site for different levels of exposure: low, medium and high.***

We conduct an experiment to test these three hypotheses. Specifically, we investigate whether thin slice judgments made about a website and its navigability by online consumers are accurate and whether they remain so for different levels of exposures. From our preceding discussion, we expect that thin slices should lead to fairly accurate judgments on the overall usability of web sites and that these judgments should remain as accurate for the smallest slice as it is for the largest slice.

Method

We plan to conduct a pilot study to verify the scales that would be used in the experiment and in the operationalization of the constructs. PONS (Profile of Non Verbal Sensitivity) scale is used to identify people who are more accurate than others in judgments. We plan to use this in our experiments to identify subjects for balanced allotment in cells. The first objective of the pretest would be to reconfirm the reliability of the Profile of Non Verbal Sensitivity (PONS) scale.

The second objective of the pilot study would be to capture the criterion ratings of the web sites that will be used for the experiment. Three experts would be exposed to twenty web sites and asked to rate the overall usability of those twenty websites in terms of the three item scale has already been tested for validity and reliability in a previous study (Agarwal and Venkatesh 2002). The three items are shown in Table 1. The experts would be given sufficient time to rate the sites on its usability to ensure objectivity and greater accuracy. We will calculate then inter-rater reliabilities to reflect the degree of agreement among the experts. Sites with the highest ratings will be chosen for the experiments. The independent variable is the exposure to the web sites and the dependent variable is their thin slice judgments. Following this, we will test whether the operationalization of the levels of exposure: low (ten seconds), medium (thirty seconds) and high (sixty seconds) works by exposing a small group of subjects to three representative web sites for the three different levels of exposure. We set the levels of exposure at 10, 30 and 60 seconds as has been used in prior research (Ambady and Rosenthal 1992; Ambady and Rosenthal 1993). Any changes to the operationalization would be carried in the final experiment. In this case, the independent variable is the exposure level to the web sites and the dependent variable is their thin slice judgments. Then, we would also examine whether the accuracy of online consumers on the navigability of a site holds for different levels of exposure. The independent variable is the exposure level to the web sites and the dependent variable is their thin slice judgments on the navigability of the site.

Our sample will consist of students because our objective is to test theory. Subjects will be chosen so that they are exposed to the Internet earlier. This will prevent any effects of learning on judgment. Subjects will not be trained as previous research has demonstrated that subjects are reliably able to make such ratings without training (Rosenthal 1987, Rosenthal et al. 1984). In order to enhance the external validity of the study, we would also ask students to play the role of consumer. This would be consistent with the methods followed in previous research (Agarwal and Venkatesh 2002). Sample size would be decided after conducting power analysis with a minimum power of 0.8, a medium effect size and an alpha level of 0.05. We approximate the sample size of each group required to test a given hypothesis at sixty students (We would use two groups to study the two hypotheses). As described earlier because some people are more accurate than others in judgments, we will identify them by PONS (Profile of Non Verbal Sensitivity) for balanced allotment in cells.

Using the guidelines provided by earlier studies, subjects would be exposed to eight sites representative of four domains: airline, bookstore, auto manufacturer and car rental. We deliberately choose these domains that have been used in a previous study (Agarwal and Venkatesh 2002) for two reasons: 1) to extend previous research and, 2) to ensure consistency between studies. The levels of exposure namely, low, medium and high will be operationalized as ten, thirty and sixty seconds following confirmation from the pilot study. In the first part of the experiment, a subject from the first sample group would be randomly exposed to 3 websites for a period of 30 seconds. After visiting each site, a subject would be asked to rate that site on its overall usability using a three item Likert scale ranging from 1 (= "extremely poor") to 7 (="extremely good"). In the second part of the experiment, each subject in the second group will be randomly exposed to a set of eight web sites for ten, thirty and sixty seconds and their judgments on the overall usability of the website and its navigability will be recorded. We will ensure that the subject is not exposed to the same site successively in order to prevent learning effects that could pose a threat to internal validity. As earlier,

after visiting each site, subjects will be asked to rate that site on its overall usability using a three item Likert scale ranging from 1 (= "extremely poor") to 7 (="extremely good").

The data analysis stage would involve calculating the accuracy between the overall usability rating of the web sites of online consumers and the rating collected from the criterion experts in the pilot study. If the correlation between the ratings is high, it would mean the accuracy of the overall usability is high.

Contribution

The objective of this study is to provide an alternate explanation of consumers' online behavior using thin slices. The findings could reconfirm the importance of thin slice judgments in an online environment and its implications on online shopping. Furthermore, the findings from this study could inform managers about the importance of web site design, including its navigability and organization. As stated earlier, consumers spend approximately fifty five seconds per page view and this number might go down in the coming months as the number of web sites keeps increasing. Therefore, it is important for managers to design their websites such that the thin slice judgments of consumers on the overall usability of a web site are favorable. Favorable thin slices arising from a given website design could turn consumers into repeat customers with long term implications for profitability and revenues.

References

Agarwal, R., and Venkatesh, V. "Assessing a Firm's Web Presence: A Heuristic Evaluation Procedure for the Measurement of Usability," *Information Systems Research* (13:2), 2002, pp. 168-186.

Ambady, N. "Telling Teachers: Predicting Student Achievement From Thin Slice Judgments," *Manuscript in Preparation*, 1999.

Ambady, N., Bernieri, F., and Richeson, J.A. "Toward a histology of social behavior: Judgmental Accuracy from Thin Slices of the Behavioral Stream," *Advances in Experimental Social Psychology* (32), 2000b, pp. 200-271.

Ambady, N., Conroy, N. M., Tobia, A., and Mullins, J. "Friends, Lovers And Strangers: Judging Dyadic Relationships From Thin Slices," *Manuscript under revision*, 2000a.

Ambady, N., Hallahan, M., and Conner, B. "Accuracy of Judgments of Sexual Orientation from Thin Slices of Behavior," *Journal of Personality and Social Psychology* (77:3), 1999, pp. 538-547.

Ambady, N., and Rosenthal. R. "Thin Slices of Expressive Behavior As Predictors Of Interpersonal Consequences: A Meta Analysis," *Psychological Bulletin* (111:2), 1992, pp. 256-274.

Ambady, N., and Rosenthal. R. "Half a Minute: Predicting Teacher Evaluations from Thin Slices of Nonverbal Behavior and Physical Attractiveness," *Journal of Personality and Social Psychology* (64:3), 1993, pp. 431-441.

Anderson, C.A., Krull, D.S., and Weiner, B. *Explanations and Processes.* In E.T. Higgins and A.W. Kruglanski (Eds), Social Psychology: Handbook of Basic Principles, New York: Guilford, 1996, pp. 271-296.

Babad, E., Bernieri, F., and Rosenthal, R. "When Less Is More Informative: Diagnosing Teacher Expectancies From Brief Samples Of Behavior," *British Journal of Educational Psychology* (59), 1989, pp. 281-295.

Babad, E., Bernieri, F., and Rosenthal, R. "Students As Judges Of Teachers' Verbal And Non Verbal Behavior," *American Educational Research Journal* (28), 1991, 211-234.

Bellman, S., Lohse, G.L., and Johnson, E.J. "Predictors of Online Buying Behavior," *Communications of the ACM* (42:12), 1999, pp. 32-38.

Bernieri, F., Sharpe, T., and Knee, C. "Social Perceptions Of Biological Motion," *Presented at the Western Psychological Association Convention*, Portland, OR, 1992.

Blanck, P.D., and Rosenthal, R. "Mediation Of Interpersonal Expectancy Effects: Counselors Tone Of Voice," *Journal of Educational Psychology* (76), 1984, pp. 418-426.

Borkenau, P., and Liebler, A. "Trait Inferences: Sources Of Validity At Zero Acquaintance," *Journal of Personality and Social Psychology* (62), 1992, pp. 645-657.

Brown, R. *Social Psychology*, New York: Free Press, 1965.

Chau, P. Y.K., Cole, M., Massey, A.P., Montoya-Weiss, M., and O'Keefe, R.M. "Cultural Differences in the Online Behavior of Consumers," *Communications of the ACM* (45:10), 2002, pp. 138-143.

Cheek, J.M. "Aggregation, Moderator Variables, And The Validity Of Personality Tests: A Peer Rating Study," *Journal of Personality and Social Psychology* (43), 1982, pp. 1254-1269.

Cutting, J.E., and Kozlowski, L.T. "Recognizing Friends By Their Walk: Gait Perception Without Familiarity Clues," *Bulletin of the Psychonomics Society* (9), 1977, pp. 353-356.

Department of Commerce. "US Department of Commerce News," **http://www.census.gov**, Retrieved from the web March 1st, 2003.

Frable, D.E.S. "Sex-Typed Education And The Perception Of Expressive Movement," *Journal of Personality and Social Psychology* (53), 1987, pp. 391-396.

Funder, D.C., and Sneed, C.D. "Behavioral Manifestations Of Personality: An Ecological Approach To Judgmental Accuracy," *Journal of Personality and Social Psychology* (64), 1993, pp. 479-490.

Gilbert, D.T., and Krull, D.S. "Seeing Less And Knowing More: The Benefits Of Perceptual Ignorance," *Journal of Personality and Social Psychology* (54), 1988, pp. 193-202.

Grudin, J. "Utility and Usability: Research Issues and Development Contexts," *Interacting with Computers* (4:2), 1992.

Hall, J.A., Roter, D.L., and Rand, C.S. "Communication Of Affect Between Patient And Physician," *Journal of Health and Social Behavior* (22), 1981, pp. 18-30.

Hecht, M.A. and LaFrance, M. "How Fast Can I Help You? Tone Of Voice And Telephone Operator Efficiency In Interactions," *Journal of Applied Social Psychology* (25), 1995, pp. 2086-2098.

Hilbert, D.M., and Redmiles, D.F. "Extracting Usability Information from User Interface Events," *ACM Computing Surveys* (32:4), 2000, pp. 384-421.

Karat, J. "Evolving the scope of user centered design," *Communications of the ACM* (40), 1997, pp. 33-38.

Kenny, D.A. (1994), *Interpersonal perception: A Social Relations Analysis*, New York: Guilford.

Klein, L.R. "Evaluating the Potential of Interactive media through a new lens: Search versus experience goods," *Journal of Business Research* (41), 1998, pp. 195-203.

Lecerof, A., and Paterno, F. "Automatic support for usability evaluation," *IEEE Transactions on software Engineering* (24), 1998, pp. 863-887.

Lippa. R. "The Naïve Research Of Masculinity-Feminity On The Basis Of Expressive Cues," *Journal of Research in Personality*, 12, 1978, pp. 1-14.

Milmoe, S., Rosenthal, R., Blane, H.T., Chafetz, M.E., and Wolf, I. "The Doctors Voice: Postdictor Of Successful Referral Of Alcoholic Patients," *Journal of Abnormal Psychology* (72), 1967, pp. 78-84.

Nielsen, J. *Usability Engineering*, Academic Press/AP Professional, Cambridge, MA, 1993.

Nielsen, J. *Designing Web Usability*, New Riders Publishing, Indianapolis, IN, 2000.

Nielsen-Netratings. "Global Internet Usage," **http://www.nielsen-netratings.com/**, Retrieved from the web March 12th, 2003.

OCLC. "Size and Growth of the Internet," **http://wcp.oclc.org/stats/size.html**, Retrieved from the web March 12th, 2003.

Palmer, J.W. "Web Site Usability, Design and Performance Metrics," *Information Systems Research* (13:2), 2002, pp. 151-167.

Prickett, T., Gada-Jain, N., and Bernieri, F.J. "First impression formation in a job interview: The first 20 seconds," *Manuscript in preparation*, University of Toledo, 1999.

Trope, Y. "Identification And Inferential Processes In Dispositional Attribution," *Psychological Review* (42), 1986, pp. 43-77.

Rosenthal, R. "Pygmalion Effects: Existence, Magnitude And Social Importance," *Educational Researcher* (16), 1987, pp. 37-41.

Rosenthal, R., Blanck, P.D., and Vannicelli. M. "Speaking To And About Patients: Predicting Therapists Tone Of Voice," *Journal of Consulting and Clinical Psychology* (52), 1984, pp. 679-686.

Shneiderman, B. *Designing the User Interface: Strategies for Effective Human Computer Interaction,* Addison-Wesley, Reading, MA, 1998.

Stallings, W.M., and Spencer, R.E. "Rating Of Instructors In Accountancy 101 From Videotape Clips," *Research report No. 265 Office of Instructional Resources: Measurement and Research Division*, University of Illinois, 1967.

Van Slyke, C., Comunale, C.L., and Belanger, F. "Gender Differences in Perceptions of Web Based shopping," *Communications of the ACM* (45:7), 2002, pp. 82-86.