

# Discrimination in reference services: A critical review of “Are virtual reference services color blind?”

Steven Chabot

Locker 90

FISI 240 Research Methods

Juris Dilevko

TA: Jane Morris

February 12, 2007

## Introduction

Shachaf and Horowitz (2006) present in this study an examination of academic libraries' virtual reference when serving culturally diversity patrons. The unobtrusive study was conducted by sending various sets of queries by e-mail to a number of Association of Research Libraries (ARL) reference services which differed only in the implicit ethnicity of the addressers' names. From their findings the writers conclude that in all dimensions Arabs and African Americans received a poorer level of service when compared with Caucasians. While these conclusions may be valid, this examination of the authors' methods and interpretations of the data will illustrate possible entry points for error and rival explanations. Given these possibilities we find that the authors' stated conclusions are overly strong and ultimately unsupported.

## Framework of the Study

The authors' stated purpose in conducting this study is to answer the question “Are virtual reference services color and gender blind and are they providing unbiased services to diverse user groups” or, more specifically, “who is more likely to be discriminated against online by virtual reference librarians?” (Shachaf & Horowitz, 2006, p. 503). This statement presents us with two problems. It implies that the authors expect some level of discrimination against minorities; such a statement suggests that the authors are approaching their research from what Babbie (2001) calls a “conflict paradigm” (p. 45). Such a theoretical framework could influence the interpretations of the data collected. It is important to note that this is not a necessary conclusion: in a follow-up study Shachaf and Oltmann (2007) found that there was no statistically significant level of discrimination in the reference service of public libraries. Secondly, the problem statement mentions gender as a possible source of discrimination, but as we will see the study does not measure gender as a variable, nor does it control for its effects.

### *Variables and Definitions*

The variables Shachaf and Horowitz intend to measure are the effects of perceived ethnicity or religion on virtual reference service quality. The authors did not give an explicit conceptual definition of what constitutes ethnicity and religion. Despite being everyday terms there is much debate surrounding their exact definition. It is beyond the scope of this review to survey all the research examining ethnicity, but recent scholarship continues to question its definition (Chandra, 2006), and Shachaf and Horowitz do not give a clear source for their classifications of ethnicity. Similarly, service quality is not explicitly defined, but we can infer from the study that has something to do with "levels of service," or the successful conformity to certain industry standards, opposed to a concept like "patron satisfaction." In this case, because the authors did not explicitly conceptually define what service quality is, we cannot judge whether the operational definition later used is adequate to this conception.

Virtual reference in this case is limited in scope to e-mail reference, based on an earlier study by Stacy-Bates (2003). Shachaf and Horowitz' study does not define the term "librarian," but if we assume this to mean a library professional who has received a MLIS or equivalent degree we will have to consider the possibility that other non-professionals also responded to the e-mail queries.

### *Literature Review*

Shachaf and Horowitz rightly identify a lack of research concerning virtual reference and cross-cultural groups (p. 501), making it slightly difficult to situate this study amongst others, although the authors do cite an article examining face-to-face reference and international students (Curry & Copeman, 2005). Curry and Copeman (2005) did conclude that patrons were receiving substandard levels of service, but did not necessarily attribute this to discriminatory bias. While Shachaf and Horowitz do cite previous studies of virtual reference service in general, the results of these studies were not presented as context for the authors' own findings: did earlier examinations of service parallel those found in the current study? We know that multiple previous studies of reference services in general have reported very high rates of patron dissatisfaction and low accuracy rates (Dewdney & Ross, 1994; Hernon & R. McClure, 1986). We cannot therefore judge the generality of Shachaf and Horowitz' findings in the context of the discipline of Library and Information Science.

However the authors do identify some key articles discussing subjective bias online from other disciplines (Douglas & McGarty, 2001; Glasser & Kahn, 2005) and there are further interdisciplinary articles, particularly from psychology, which support this study. For instance, one study found that subjects were more likely to express discriminatory biases when interviews were conducted over the Internet opposed to in the lab (Evans, Garcia, Garcia, & Baron, 2003). In support of the authors' use of ethnic names to measure biases, studies of fictitious replies to rental ads (Carpusor & Loges, 2006) and help-wanted ads (Bertrand & Mullainathan, 2004) confirm Shachaf and Horowitz's conclusions that ethnic names illicit discrimination. Furthermore, this discrimination occurs not only when ethnicity is perceived in proper names, but also when it is inferred from e-mail addresses themselves (Pittinsky & Shih, 2006).

In reference to the support for their investigative method in the literature, Shachaf and Horowitz do

acknowledge that other researchers have chosen different methods to evaluate reference services, one method in particular being surveys of user satisfaction (Hernon & Calvert, 2005). Nevertheless, Shachaf and Horowitz do make a convincing case for the use of unobtrusive methods in this study (p. 504), given this method's long use in evaluating reference services. Also, the fact that library staff would not be under direct observation means they would be more likely to exhibit natural reactions when receiving the test e-mails, a supposition supported by Evans (2003).

## Methods

The methods of this study specifically measured a certain set of criteria coded from librarian's replies to fake e-mail queries posed to 23 ARL member libraries' virtual reference services. To control for possible sources of error the content of e-mails were drawn from six pre-composed queries with only details changed for each particular university ("Could you tell me the population of [institution's city name] in 1963 and 1993?" (p. 505)) and the false name attached. The study was not repeated beyond the one test group.

Different names were attached to each e-mail through which the ethnicity and religion were to be implicitly judged by the reference staff providing service. However, as we have noted, the judging of names can be ambiguous, and the categories of ethnicity were not rigorously defined.. For instance, given the definition by the U.S. Census Bureau (2001), *Mary Anderson*, *Ahmed Ibrahim* and *Moshe Cohen* would all be considered White. In this study *Mary Anderson* was considered the sole Caucasian at one point (p. 505), but this categorization included *Moshe* at another point (p. 512). Furthermore, Shachaf and Horowitz do not sufficiently discuss the distinction between Judaism as an ethnicity or as a religion (Yerushalmi, 1991; Webber, 1997). *Ahmed* himself is sometimes Muslim, a designation of religion (p. 503) and at other times an Arab, a designation of ethnicity (p. 508).

As well, the study was deficient in the fact that it attempted to measure various possible sources of discrimination at once. We noted above that the preliminary problem statement included gender as a variable to be studied although none of the later interpretations take gender into account. And, as we have indicated, there is no rigorous distinction made throughout the study between ethnicity and religion. To guard against possible rival explanations only one of the three should have been measured at once, as was done in Shachaf and Oltmann's (2007) follow-up study. As such, we can never be entirely sure if the discrimination found in this study is a factor of gender, ethnicity or religion.

In operationalizing service quality, the authors used the guidelines on virtual reference published by the International Federation of Library Associations and Institutions (IFLA) (2006) and the Reference and User Services Association (RUSA) (2007) as standards by which to judge the library staff's level of service. These are adequate definitions given that the libraries themselves have consented to adopting them as guidelines. Level of service is judged based on 23 measurements including but not limited to the word length of replies, elapsed time from message sent to first response, and the measurement of certain textual contents such as the forms of greetings and signatures. All of these measurements were collected and averaged, but the authors caution that not enough samples were taken for statistical generalizations (p. 508, 517-8). The tallying of these measurements is relatively straightforward, and the use of two coders gives us a level of confidence in their accuracy.

However, as noted above the authors did not control for the identity of the library staff who answered the e-mails. These may often have been student library assistants who may or may not be expected to be knowledgeable about best practices and industry standards in this case. This means, therefore, that the methods of this study can not accurately measure discrimination in librarians, but at best can only measure the varying levels of service in reference departments as a whole.

A final source of error which occurred during the measurement phase of the study is the fact that two of the batches of e-mails, sent under *Mary* and *Latoya*, were sent on the weekend, calling into question the calculations of response time for these two names, and possibly some of the other measurements as well.

### Findings and Summaries

The Shachaf and Horowitz conclude that "on all the dimensions of service quality that were evaluated" Arabs (Muslims?) and African Americans were discriminated against (p. 512). Apart from the possible sources of error we have already suggested, we feel that this conclusion is overly strong given some of the results of the study. Often these two names were given average or above-average service compared with the other names. For instance, in the measure of total answers (Fig. 5) *Ahmed* received the median number, and in the measure of follow-ups and concluding remarks (Fig. 9) he was equal with the Caucasian name, *Mary*, although we concede that he did receive a poor level of service on such measurements as response time (Fig. 1). As well, some level of bias was exhibited through the authors' dismissal of one measurement where *Ahmed* received the highest level of service as "anecdotal" (p. 516). Looking at *Moshe*, he seems to have overly inflated results under both response time and length of reply (Fig. 3), making us wonder if he received a response which distorted his averages. The most we could conclude from these findings is that on some of the measurements there was a level of discrimination, but more study is required for solid conclusions.

Particularly troubling were the implications drawn by the authors from these interpretations: "Findings indicate that librarians respected Moshe and Mary more than they respected Ahmed" (p. 514). This conclusion goes beyond the scope of the study, which was designed to measure service quality, not levels of respect. Shachaf and Horowitz admit that anxiety about names could account for some of the low scores for both *Ahmed* and *Chang Su* (p. 514). While we do not deny that some of the measures could indicate discrimination, one cannot conclude lack of respect from these findings alone.

### Conclusion

Given the ambiguous definitions, the lack of control for gender when measuring ethnicity, and the other sources of error we have identified, we conclude that the results of this study only suggest the need for further investigation. While we agree with Shachaf and Horowitz in admitting the results of this study cannot be generalized, we disagree with the strength of the authors' conclusions. At best the study indicates some level of possible discrimination against these groups, but the data is far from conclusive on all points, and, given the level of possible error, such a conclusion can only be made tentatively. Furthermore, while the literature in this area does support the authors' hypothesis, given

another set of libraries at a different time, we would not be surprised to see the levels of service differ substantially. With these considerations, we cannot affirm the suggested causal connection between ethnicity and virtual reference service levels at this time without further research.

## References

- Babbie, E. R. (2001). *The practice of social research* (9th ed.). Stamford, CT: Wadsworth.
- Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *The American Economic Review*, 94(4), 991-1013.
- Carpusor, A. G., & Loges, W. E. (2006). Rental discrimination and ethnicity in names. *Journal of Applied Social Psychology*, 36(4), 934-952.
- Chandra, K. (2006). What is ethnic identity and does it matter? *Annual Review of Political Science*, 9, 397-424.
- Curry, A., & Copeman, D. (2005). Reference service to international students: A field stimulation research study. *The Journal of Academic Librarianship*, 31(5), 409-420.
- Dewdney, P., & Ross, C. S. (1994). Flying a light aircraft: Reference service evaluation from a user's viewpoint. *RQ*, 34(2), 217-231.
- Douglas, K. M., & McGarty, C. (2001). Identifiability and self-presentation: Computer-mediated communication and intergroup interaction. *British Journal of Social Psychology*, 40(3), 399-416.
- Evans, D. C., Garcia, D. J., Garcia, D. M., & Baron, R. S. (2003). In the privacy of their own homes: Using the internet to assess racial bias. *Personality and Social Psychology Bulletin*, 29(2), 273-284.
- Glasser, J., & Kahn, K. B. (2005). Prejudice and discrimination and the Internet. In Y. Amichai-Hamburger (Ed.), *The social psychology of the Internet* (p. 247-274). Oxford: University Press.
- Hernon, P., & Calvert, P. (2005). E-service quality in libraries: Exploring its features and dimensions. *Library & Information Science Research*, 27, 377-404.
- Hernon, P., & McClure, C. (1986, April 15). Unobtrusive reference testing: The 55 percent rule. *Library Journal*, 111(7), 37-42.
- International Federation of Library Associations and Institutions. (2006, December). *IFLA digital reference guidelines*. Retrieved February 11, 2007 from IFLANET: <http://www.ifla.org/VII/s36/pubs/drgo3.htm>.
- Pittinsky, T. L., & Shih, M. J. (2006). Identity cues: Evidence from and for intra-individual perspectives on positive and negative stereotyping. *Journal of Applied Social Psychology*, 36(9), 2215-2239.
- Reference and User Services Association. (2007, January). *Guidelines for implementing and maintaining virtual reference services*. Retrieved February 11, 2007 from: <http://www.ala.org/ala/rusa/rusaprotools/referenceguide/virtrefguidelines.htm>.
- Shachaf, P., & Horowitz, S. (2006). Are virtual reference services color blind? *Library & Information Science Research*, 28(4), 501-520.
- Shachaf, P., & Oltmann, S. (2007). E-quality and e-service equality. In *Proceedings of the fortieth Hawaii*

*international conference on systems sciences*. Los Alamitos, CA: IEEE Press. Retrieved February 11, 2007 from IEEE Computer Society Digital Library:

<http://csdl.computer.org/comp/proceedings/hicss/2007/2755/00/27550247c.pdf>.

Stacy-Bates, K. (2003). E-mail reference responses from academic ARL libraries: An unobtrusive study. *Reference & User Service Quarterly*, 43(1), 59-70.

United States Census Bureau. (2001). *Overview of Race and Hispanic Origin*. Retrieved February 11, 2007 from <http://www.census.gov/prod/2001pubs/c2kbr01-1.pdf>.

Webber, J. (1997). Jews and Judaism in contemporary Europe: Religion or ethnic group? *Ethnic and Racial Studies*, 20(2), 257-279.

Yerushalmi, Y. H. (1991). *Freud's Moses: Judaism terminable and interminable*. New Haven: Yale.