Variation in Country-Based Ranking Lists Among Consumers’ Choices of Top E-Commerce Web Sites: Implications for International E-Marketing

Bruce W.N. Lo  
Department of Information Systems  
University of Wisconsin-Eau Claire, Eau Claire WI 54702, USA  
lobw@uwec.edu

Faye J. Kao  
Department of Marketing  
Eastern Michigan University, Ypsilanti, MI 48197, USA  
fkao@emich.edu

ABSTRACT
To develop an understanding of the factors that may impact on international e-business and e-marketing, this research examines the top 100 e-commerce Web sites in a sample of nine countries of diverse cultures and at different stages of national development, to determine the extent to which the forces of localization and globalization are influencing consumers’ preference and choices in ranking e-commerce Web sites in these countries. Preliminary results show evidence of variation among these ranking lists, presumably because of regional cultural preferences, local demands, and language differences. There is also evidence of some common characteristics among these lists because of the globalizing influence of the Internet. The implications of these findings for international e-marketing are discussed.

Keywords: e-commerce, e-marketing, Web site ranking, top-100 Web sites, localization, globalization, culture
1. INTRODUCTION

The number of Internet users has increased dramatically in the last few years. Table 1, from the Internet World Statistics Web site (Internetworldstats 2008), shows that the world’s online population experienced a spectacular, near-300% growth during the period 2000–2007. The largest growths are outside North America – in the Middle East, Africa, Latin America, and Asia. In fact, the largest online group, Asia, represents nearly 38% of the world total. The Internet “penetration” data indicate that the greatest potential for growth lies outside the developed nations. Therefore, any e-businesses operating on the Internet have a growing international presence and cannot ignore their global marketing image. The question of how local culture and global influence of e-commerce may shape customers’ Web site choice behavior is something that requires careful attention.

<table>
<thead>
<tr>
<th>World Region</th>
<th>Population</th>
<th>% World Population</th>
<th>Number of Internet Users</th>
<th>Penetration (% of Pop.)</th>
<th>% World Use</th>
<th>% Growth (2000-2007)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Africa</td>
<td>955,206,348</td>
<td>14.3</td>
<td>51,022,400</td>
<td>5.3</td>
<td>3.6</td>
<td>1,030.2</td>
</tr>
<tr>
<td>Asia</td>
<td>3,776,181,949</td>
<td>56.6</td>
<td>530,153,451</td>
<td>14.0</td>
<td>37.5</td>
<td>363.4</td>
</tr>
<tr>
<td>Europe</td>
<td>800,401,065</td>
<td>12.0</td>
<td>384,332,394</td>
<td>48.0</td>
<td>27.2</td>
<td>265.7</td>
</tr>
<tr>
<td>Middle East</td>
<td>197,090,443</td>
<td>3.0</td>
<td>41,939,200</td>
<td>21.3</td>
<td>3.0</td>
<td>1,176.8</td>
</tr>
<tr>
<td>North America</td>
<td>337,167,248</td>
<td>5.1</td>
<td>247,637,606</td>
<td>73.4</td>
<td>17.5</td>
<td>129.1</td>
</tr>
<tr>
<td>Latin America</td>
<td>576,091,673</td>
<td>8.6</td>
<td>137,200,309</td>
<td>23.8</td>
<td>9.7</td>
<td>659.3</td>
</tr>
<tr>
<td>Australasia / Pacific</td>
<td>33,981,562</td>
<td>0.5</td>
<td>20,204,292</td>
<td>59.5</td>
<td>1.4</td>
<td>165.1</td>
</tr>
<tr>
<td>WORLD TOTAL</td>
<td>6,676,120,288</td>
<td>100</td>
<td>1,412,489,652</td>
<td>21.2</td>
<td>100</td>
<td>291.3</td>
</tr>
</tbody>
</table>
Because of its global presence, e-commerce provides a fruitful area to study a phenomenon that has attracted a great deal of interest among marketing researchers. The aim of this research is to investigate the relationship between local or national culture and the global influence of e-commerce in cyberspace. In particular, the present paper focuses on examining whether the top-ranking Web sites from countries of diverse culture and at different stages of their national development exhibit any different pattern in consumer preferences that may be attributed to the influence of local cultures.

In a recent paper, Lo and Martinez Santa Cruz [2007] examined the localization effect from a very different point of view. There, Web site lists were chosen from specific “ethno-linguistic” groups where the Web sites originated. This represents a “provider” perspective for the Web site ranking lists, because the lists are based on the language used by the Web site providers.

In the current study, we chose the website ranking lists based on the consumers’ choice from all available e-business sites, regardless of which language the sites are in and from which country they originate. Therefore, this is a “consumer” perspective in selecting the Web site ranking lists. Furthermore, the previous paper examined the “contents” of the sites, whereas this study focuses on the similarities/differences of the ranking lists themselves and not on the content or function of the sites. It is our belief that this approach may provide a rich source of information for the current investigation.

The purpose of this study, therefore, is to seek clarification of the effect of localization on international e-commerce by examining the leading Web sites from different countries, from the consumer’s angle, to see whether there is evidence of cultural variation and/or globalization influence.

The specific research questions that will be considered in this study are:

- Do the top-ranking Web site lists from different countries exhibit any different characteristics?
  - common characteristics?
- If there are differences, can these be related to local or regional culture?
- If there are commonalities, can they be attributed to the globalization influence in cyberspace?
- Do these differences/commonalities vary according to the size of the ranking list? If so, what does this mean?
- What lessons can one learn from these observations in relation to international marketing for e-business?

The remainder of this paper is organized into four sections. The next section, Section 2, provides a background and rationale for this study by examining the body of relevant literatures. This is followed by Section 3, which describes the methods used in this research. It describes both the source of data and the computational procedures used in arriving at the results. Section 4 presents the results, and Section 5 discusses the results and their implications. The paper concludes with comments on possible future research directions.
2. BACKGROUND OF STUDY

Culture is an important consideration in international business and particularly in international e-business. Leung, Bhagat, Buchan, Erez and Gibson [2005] broadly define national culture as the values, beliefs, norms, and behavioral patterns of a national group, which they believe will influence the decision and behavior pattern of consumers in different parts of the world as the groups engage in international e-business. Traditionally, marketing research has adapted overwhelmingly a dichotomization approach leading to a “standardization versus localization” debate [Joy and Wallendorf, 1996; Bin, Chen, and Sun, 2003; Singh, Furrer, and Ostinelli, 2004]. The main assumption is that in a less developed economy, consumers would adopt one of two mutually exclusive positions – either they would remain traditional, preferring advertisements that match local culture, or they would embrace the dominant culture of the day, preferring advertisements that promote global culture. The former would favor a local marketing strategy, whereas the latter would favor a global marketing strategy. However, the findings of more recent studies are mixed [Chang, Wong, and Koh, 2003; Zhou and Belk, 2004], and the reality is more complex than this simple dichotomy.

Globalization brings different cultures into close interaction. The result is that both culture homogenization and cultural fragmentation are occurring at the same time. On the one hand, because of the presence of global brands, there emerges a global consumer culture, wherein consumers from different countries share the same consumption meaning, value, and behavior [Alden, Steenkamp, and Batra, 1999]. On the other hand, localization occurs as soon as foreign consumer objects are brought in for local consumption [Hannerz, 1989; Yoshimoto, 1989; Chau, Cole, Massey, Montoya-Weiss, and O’Keefe, 2002]. The intersection of global commodities and local consumption culture results in culture fragmentation. The fragmentation manifests in music, food, media content, etc. [Appadurai, 1990].

With respect to the globalization influence of e-commerce, there are two opposing points of view among scholars. Some argue that the dominance of transnational e-commerce giants have created a converging trend of standardization in Web communication and Web contents [Singh, Zhao, and Hu, 2003; Singh, Furrer, and Ostinelli, 2004]. Others argue that there is evidence of cultural sensitivity in thematic appearance, contents, interactivity, and consumer trust behavior [Sagi, Carayannis, Dasgupta, and Thomas, 2004; Lo and Gong, 2005; and Wurtz, 2005]. Therefore, Web site localization is critical to business success. Although some recent studies appear to lean toward the latter, the issue is by no means closed.
This is not too dissimilar to the dichotomized debate in international advertising where, on the one hand, marketing scholars argue that, under the standardization approach, the globalization process will absorb different national cultures and meld them into a global consumer culture. On the other hand, some scholars argue that there is clear evidence that localization forces shaped by national/regional culture are at work in international marketing, thus resulting in a diverging influence [Joy and Wallendorf, 1996]. More recently, scholars began to realize that consumer preference may be a bit more complex than a simple dichotomy. In a transitional economy, consumers may apply a variety of approaches, such as different cultural interpretive strategies and self-referencing, to absorb global images into their cultural and consumption schema [Hung, Li, and Belk, 2007]. Therefore, homogenization and heterogenization may occur at the same time. This is sometimes referred to as *glocalization*.

Researchers have now come to the realization that traditionalism and modernity are not mutually exclusive [Smith and Bond, 1998]. The globalization of production, capital, and the media only leads to “partial” globalization of consumer culture. In fact, globalizing forces are reinvigorating traditional culture in many regions of the world, where there is a solid history of national cultural characteristics. Instead of moving monotonically toward globalization, a national culture converges and diverges to redefine what is meaningful to that nation and its people for that given time [Leung et al., 2005]. The forces at work are not just unidirectional, but are also multidirectional. Several interesting studies [Zhang, Zheng, and Wang, 2003; Chang, Wong, and Koh, 2003] found that Chinese people endorse both traditional and modern values. Therefore, both values coexist in contemporary Chinese advertising [Zhang and Shavitt, 2003; Zhou and Belk, 2004].

Although the studies mentioned above were conducted mainly in a single country, we believe that this phenomenon of glocalization is applicable to other national groups. This leads us to the current study. Furthermore, instead of looking at traditional international marketing activities, we propose to examine Web-based international e-business. One way to study this is to examine how consumer groups in different countries view the e-business Web sites on the Internet. Do they prefer to visit those e-business sites that reflect their own traditional and/or unique cultural values or do they go for the global e-commerce giants whose brands they recognize?

Presumably, their Web-site-browsing behavior patterns will tell us something about their preferences. Therefore, it is proposed that we examine the top-ranking e-business Web site lists in a selection of different countries. It is hoped that the sites included in their top 100 list and the rank order in which they appear will tell us something about the consumers’ preferences. The remainder of this paper will proceed to describe how we go about comparing the top 100 e-business Web site lists from different countries.
3. METHODS

The main task of this study involves the comparison of Web site ranking lists from different countries to discover commonalities and differences among the lists. This section will describe which lists were chosen for our analysis and how we performed the comparison.

3.1. Source of Data

There are many ranking list providers for e-business Web sites, using a variety of methods to rank the sites. Table 2 provides a sample of different ranking providers.

<table>
<thead>
<tr>
<th>Table 2. Some Web Site Ranking Providers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alexa Internet, Inc. <a href="http://www.alexa.com">http://www.alexa.com</a>;</td>
</tr>
<tr>
<td>100BestWebsites, <a href="http://www.100bestwebsites.org">http://www.100bestwebsites.org</a>;</td>
</tr>
<tr>
<td>ComScore Media Metrix, <a href="http://www.comScore.com">http://www.comScore.com</a>;</td>
</tr>
<tr>
<td>Nielsen, <a href="http://www.nielsen-netratings.com">http://www.nielsen-netratings.com</a>;</td>
</tr>
<tr>
<td>PC Magazine, <a href="http://www.pcmag.com/article2/0,1759,1554010,00.asp">http://www.pcmag.com/article2/0,1759,1554010,00.asp</a>;</td>
</tr>
<tr>
<td>Ranking.com <a href="http://www.ranking.com">http://www.ranking.com</a>;</td>
</tr>
<tr>
<td>Time’s, <a href="http://www.time.com/time/2005/websites">http://www.time.com/time/2005/websites</a>;</td>
</tr>
<tr>
<td>Websearch, <a href="http://www.websearch.com">http://www.websearch.com</a>;</td>
</tr>
<tr>
<td>Web100, <a href="http://www.web100.com">http://www.web100.com</a>;</td>
</tr>
</tbody>
</table>

The different ranking methods basically fall into three categories, as described below. For a more complete description, see the paper by Lo and Sedheim [2006].

a. Activity-based criteria, also known as usage-based or traffic-based ranking, are probably the best known and most researched among the three. This method is usually regarded as the most objective. Here, Web sites are ranked according to the volume of activity that takes place on the site. The site that attracts the most traffic or has the highest number of user visits would rank at the top. The more discerning readers will
recognize that the concept of a “site visit” (or “hits”) is by no means simple. It may be defined in terms of a number of different meanings: *reach, frequency, duration,* or a variation on any of these or a combination thereof. The reader is referred to two earlier papers [Lee and Leckenby, 1998; and Lo and Sedheim, 2006] for a fuller discussion on these issues. Ranking providers that claim to use this criterion include Alexa, comScore Media Metris, Nielsen NetRatings, Ranking.com, and Websearch.

b. **Reference-based criteria** rank a Web site according to how frequently that site was cited by another in relation to a given search topic. Presumably, the more frequently a site is cited by other sites, particularly sites that are regarded as authorities on the subject matter, the more important that link will be weighted. The Google page-rank method is probably the best known among this group of methods.

c. **Opinion-based criteria** use the opinion of a panel of judges to rank the list of Web sites. The resulting rankings thus depend on the subjective judgment and preferences of the panel members. In a sense, all ranking methods, including the traffic-based ones, rely on a panel of judges because, in the traffic-based criteria, an Internet surfer decides to visit a Web site or not. But what distinguishes opinion-based criteria from the others is their explicit dependence on the subjective opinions of the judges, with little or no regard for objective statistics. It is not surprising to find that opinion-based ranking lists may be very different from traffic-based lists. Among ranking providers cited above, the following may be classified in this category: *Time*’s 50 coolest Web sites, *PC Magazine*’s top 100 Web sites, 100 Best Web sites, World’s hottest sites, and Web100.

In this study we base our analyses on ranking lists provided by www.Alexa.com [Alexa, 2007]. Alexa is chosen for several reasons. It is well-known and uses the more reliable “traffic-based” method for Web site ranking. Alexa’s parent company is Amazon Web Service, which is a well-known and reliable company. Alexa data are free, and are widely available. The Alexa toolbar has a large installed base of users. The Alexa data are constantly updated.

Once it has been decided which ranking provider to use, the next question is: Which countries’ site ranking data are we going to collect? Remembering that our objective is to examine differences and commonalities among these Web site ranking lists, it would make sense to include a collection of countries that have varied cultural backgrounds and are at different stages of national development. In this study, we included the following nine countries:
Variation in Country-Based Ranking Lists Among Consumers’ Choices of Top E-Commerce Web Sites

The USA, France, and Germany are chosen to be representative of the Western culture, but with different languages. Indonesia and Pakistan are chosen because of their Islamic heritage. India and Pakistan are chosen because of their common Indo-subcontinent culture and because both are strongly influenced by Britain and the English language. Mexico represents the Latin American group, whereas China and Japan represent the East Asian culture. The ranking lists were collected around September 7, 2007. It is recognized that Web site rankings are not static and may change. For the current analysis, we take a snapshot. The phenomenon of ranking changes over time will be the subject of a future investigation.

In addition to the nine lists of top-100 e-business Web sites in these nine countries, we also collected the data for the “global” top-100 Web sites. By global, we mean the top-100 most visited Web sites overall, regardless of from which country the consumers come. The “global” top-100 list will be used as the “reference” from which each country list may be compared. To save space, Table 3 shows just the top-10 Web sites for the ten lists mentioned above (nine countries, plus global). The full top-100 lists are available from the authors by request. In Table 3, GL refers to the “global list,” whereas the standard country domain codes are used for each of the nine countries.

3.2. Computational Details

To investigate the differences (or similarities) between two rank-order lists, we propose a three-stage approach similar to that used in a previous paper to compare ranking providers [Lo and Sedheim, 2006]. These three stages are progressively more discriminating in characterizing the differences.

(1) Membership Commonality (m) – Here, we compute the percentage of common membership between two lists of the same size. For example, the question we ask is this: Among the top-20 Web sites from the two countries, how many of them are common to both lists? Suppose 10 out of the top-20 sites from one country are also among the top-20 sites in another country. The value of m=50% or 0.5; that is, half of the Web sites in each list are the same. The higher this number, the more commonality between the two lists. If m=1, this means the two lists are identical, implying that consumers from the two countries from which the ranking lists were based agree on which e-business sites should be included among the top-n ranking list. In other words, the consumers in

| USA | Pakistan | Indonesia |
| Mexico | France | Germany |
| India | Japan | China |
the two countries exhibit similar choice and preference behaviors with respect to e-business Web site ranking. On the other hand, if $m=0$, the two lists have nothing in common, implying that consumers from the two countries completely disagree on their choice of the top-n ranking Web sites. It should be pointed out that we did not just compute a single value of $m$. In order to observe the similarity and diversity pattern, we compute a value of $m$ for lists of different sizes $n$, where $n = 1, 2, 3, \ldots 100$; that is, an “$m$-value” for each of the “top-n” list. Therefore, we have a total of 100 $m$ values. What we did was calculate the membership commonality as a function of the size of the Web site ranking list.

(2) **Kendal’s Rank Correlation ($\tau$)** – The second metric extracts the “common” lists from the ranking lists originated from the two different countries, and then computes the Kendal’s tau, $\tau$, rank correlation
coefficient [Daniel, 1978] for the two common lists. Even though the membership of the common list from country 1 may be the same as that of country 2, the ranking of their Web sites may still differ. For example, let us consider the common lists of the U.S. and Mexico when the list size n=3. For illustration purposes, let us say that, for the U.S., the top-3 sites are Yahoo, Google, and MSN; but the top-3 sites for Mexico are Yahoo, MSN, and Google. In this case, the memberships are the same, yet the ranks of the sites are different. So we can easily compute the Kendall’s tau for the two lists with the standard formula [Wikipedia, 2007].

\[
\tau = \frac{4P}{n(n-1)} - 1
\]

where, \( P \) is the number of concordant pairs, while \( n \) is the number of items in the two lists.

In their paper, Lo and Sedheim [2006] also compute the Spearman’s rho rank correlation coefficient, but no new information was gained by that computation. So we shall only compute tau (\( \tau \)) and ignore rho in the present study. Again, we did not compute just one value of tau (\( \tau \)). We in fact computed a tau value for each pair of lists of size \( n=1, 2, 3, \ldots, n_{maxcom} \), where \( n_{maxcom} \) is the maximum number of common members between the two lists from the two countries. For example, for the US top-100 list and the Mexico top-100 list, \( n_{maxcom} = 39 \); that is, they contain 39 websites common to both lists. To recap, we note that the membership commonality (\( m \)) simply measures whether two lists contain the same set of Web sites, but the Kendall’s tau (\( \tau \)), which is more discriminating, measures to what degree the rank orders of the elements in the two lists are similar.

(3) **Rank Discordance Index (\( \Delta \))** - The final metric delta (\( \Delta \)) measures the amount of discordance between the two common lists. It is calculated as a normalized value of the actual sum of squares of the rank differences between the two common lists. This is a new measure that was introduced by Lo and Sedheim [2006]. Readers are referred to that paper for computational details. For ease of reference, we shall call this the “Rank Discordance Index” (RDI, for short), which will be denoted by the symbol delta, \( \Delta \). Compared with the two previous metrics, this is an even more discriminating measure. Tau (\( \tau \)) only compares the rank order of the lists, but delta, \( \Delta \), actually takes into consideration the rank differences between the two common lists with respect to the original rank positions of the Web sites in the country-based lists. Even though
the rank order of a pair of sites may be the same, their rank positions in the original country-based ranking lists may still differ. The value of delta, $\Delta$, for lists of size $n$, may be computed by the formula [Lo and Sedheim, 2006]:

$$
\Delta = K \left(2 - n_{\text{maxcom}}/n\right) \sum_{j=1}^{n_{\text{maxcom}}} (r_{j1} - r_{j2})^2 / [n(n^2-1)]
$$

where, $K$ is a constant, and $n_{\text{maxcom}}$ is the number of common members between the two lists of size $n$; $r_{j1}$ is the rank of site $j$ in the original country 1 list, whereas $r_{j2}$ is the rank of site $j$ (remember this site is common to both lists) in the original country 2 list. The actual value that was calculated is, in fact, $\Delta / K$ and not just the delta, since we do not really know the exact value of the constant $K$. Again, instead of computing just one value of delta ($\Delta$), we compute a $\Delta$ value for each pair of lists of size $n=1, 2, 3, ..., n_{\text{maxcom}}$, where $n_{\text{maxcom}}$ is the maximum number of common members between the two country-based lists.

4. RESULTS

In this section, we present the results of our analyses. Since we performed three sets of analyses, we present the computational outcomes in three subsections.

4.1. Membership Commonality (m) Results

Table 4 gives the “membership commonality” matrix for the top-100 Web sites from the global ranking lists, plus the lists for the nine countries.

<table>
<thead>
<tr>
<th></th>
<th>Gl</th>
<th>us</th>
<th>mx</th>
<th>In</th>
<th>pk</th>
<th>de</th>
<th>fr</th>
<th>id</th>
<th>jp</th>
<th>cn</th>
</tr>
</thead>
<tbody>
<tr>
<td>gl</td>
<td>1.00</td>
<td>0.41</td>
<td>0.41</td>
<td>0.31</td>
<td>0.31</td>
<td>0.30</td>
<td>0.28</td>
<td>0.28</td>
<td>0.19</td>
<td>0.20</td>
</tr>
<tr>
<td>us</td>
<td>1.00</td>
<td>0.39</td>
<td>0.33</td>
<td>0.35</td>
<td>0.27</td>
<td>0.22</td>
<td>0.31</td>
<td>0.16</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>mx</td>
<td>1.00</td>
<td>0.26</td>
<td>0.29</td>
<td>0.27</td>
<td>0.25</td>
<td>0.29</td>
<td>0.16</td>
<td>0.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>In</td>
<td>1.00</td>
<td>0.46</td>
<td>0.20</td>
<td>0.19</td>
<td>0.26</td>
<td>0.12</td>
<td>0.10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pk</td>
<td>1.00</td>
<td>0.22</td>
<td>0.20</td>
<td>0.34</td>
<td>0.13</td>
<td>0.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>de</td>
<td>1.00</td>
<td>0.19</td>
<td>0.22</td>
<td>0.12</td>
<td>0.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fr</td>
<td>1.00</td>
<td>0.18</td>
<td>0.13</td>
<td>0.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>id</td>
<td>1.00</td>
<td>0.12</td>
<td>0.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>jp</td>
<td>1.00</td>
<td>0.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cn</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The interpretation of this table is as follows. Take, for example, the intersection of the second row (US) with the third column (MX), which contains the number 0.39. This means that, out of the top-100 website sites, the US and Mexico have 39 of them in common between the two lists.

As mentioned earlier, we computed the values of \( m \) for lists of all different sizes \( n \), where \( n \) ranges 1 to 100. This means that, for each pair of countries, we computed an \( m \) value for every \( n \) value, resulting in 100 such \( m \) values. Including the “global” list, there are a total of 10 lists. So there are \( 10 \times C_2 = 45 \) such pairs. But for each pair of countries, there are 100 points. Therefore, there are a total of 4,500 such \( m \)-values. To make them more comprehensible, we shall present the results graphically, plotting the \( m \)-value as a function of \( n \). Figure 1 shows the nine curves, pairing up the “global” list with the lists from each of the nine countries.

Note that all nine curves exhibit some common features – they peak at around \( n = 5 \) to \( n \approx 20 \), and that they all approach some asymptotic values as \( n \) gets larger. The curves also bunch into three groups: US and Mexico fall into one group approaching the value of \( m = 0.4 \) when \( n \) is large, whereas India, Pakistan, Germany, and France fall into a second group approaching the value of \( m = 0.3 \) when \( n \) is large, and Japan and China fall into a third group approaching the value of \( m = 0.2 \) when \( n \) is large. These three groups roughly correspond to what we may refer to as the “North American group,” the “Indo-European group,” and the “East Asia group,” respectively.
The most unexpected result is probably the “Indo-European group.” To see whether this observation is real, we also present the membership commonality graph using Germany (instead of GL) as a reference point for comparison with the other countries. This is shown in Figure 2.

![Figure 2: Membership Commonality DE with Others](image)

Again, this graph shows, particularly when the value of $n$ is large, the three groupings of countries; i.e., the North American group of US and Mexico; the Indo-European group of India, Pakistan, and France (Germany is not graphed here as it makes no sense to correlate with itself); and the East Asian group of Japan and China. For this reason, we feel that the observed pattern may be considered real.

A natural question is whether the second group of four countries (India, Pakistan, Germany, and France) are indeed a single grouping, or whether there are sub-groupings. The results of comparing India with the other three countries are shown in Figure 3. Not surprisingly, India and Pakistan seem to have a lot more in common with each other than with the other three countries. Note the substantially higher asymptotic $m$-value (close to 0.5) between India and Pakistan. Not so obvious is the India-Indonesia commonality curve, which seems also to have a slighter higher value than those of France or Germany. This provides a finer subdivision of India and Pakistan, versus German and France, with respect to Web site ranking patterns.

Thus far, we have been looking at the question of the degree to which the Web site ranking lists from two different countries share common memberships. As pointed out earlier, even if two lists have common memberships, the rank order of each site within each list may still differ. So, we need to perform a more discriminating analysis and determine to what degree the rank orders in the two lists are in agreement (not just the membership). This brings us to the next set of analyses.
4.2. Rank Correlation (τ) Results

We use the Kendall’s rank correlation (τ) to measure the degree of similarity in rank order between the two ranking lists from the two countries. The question we ask here is this: Suppose the ranking list from country 1 and the ranking list from country 2 have \( n_{\text{maxcom}} \) Web sites in common, how does the rank order in list 1 correlate with the rank order in list 2? In other words, what we are examining is: To what extent do the two sets of rank orders correspond to each other?

We present the results in four graphs. In each case, we correlate the “global” ranking list with a list from another country; so, there will be a total of nine curves.

In the first graph, Figure 4, we diagram the tau values for the GL-US and GL-MX correlations. The two countries, US and Mexico, are grouped together because their curves are likely to behavior rather similarly. In Figure 4, we also include the value of \( \tau^* \), which is the critical value of Kendall’s tau. If the computed tau statistic is greater than this critical value, \( \tau^* \), then we cannot rule out the fact that the two rank order lists are significantly correlated. In Figure 4, this means that, for lists larger than about 15, there appears to be a significant correlation between the US and MX ranking lists with the “global” list. In other words, when the rank order list is large enough, some of the better known Web sites begin to find their respective positions in both country-based lists.
Figure 5 presents the results for the correlation of the global list with lists from three other countries: India, Pakistan, and Indonesia. These three are grouped together because, in the previous section, there are suggestions that they may exhibit similar preference patterns. The relationship between India and Pakistan is obvious as they have a common British connection, and in fact were originally one country, before splitting up into two countries (one predominantly Islamic and the other predominantly Hindu). The relationship with Indonesia, another Islamic country, is less obvious. Perhaps the common factor here is the dominance of the English language in these countries. Figure 5 shows that all three curves cross the \( \tau^* \) (significant threshold) curve at around \( n = 15 \), as in the case of Figure 4 for the US and Mexico.
Figure 6 shows the correlation between the global list and the two ranking lists from the two European countries, Germany and France. These two curves cross the significant threshold (τ*) curve much earlier, around \( n = 10 \).

Figure 7 presents the results for Kendall’s τ for correlating the global list with that of Japan and China. The Japan and China τ curves never quite cross the critical value line \( τ^* \). This probably means that there are sufficiently strong forces at work, culturally or otherwise, so that the rank list preferences from these two countries are sufficiently different from the “global” behavior. It is probably reasonable to say that the ranking lists from China and Japan exhibit a significantly different pattern than those from the other seven countries.
4.3. Rank Discordance Index (Δ) Results

The final analysis is concerned with the discordance between rank orders among ranking lists from different countries. Here, we take into consideration the distance between the rank differences of the original lists, and not just their rank order. The computation results are presented in Figures 8, 9, and 10.

Figure 8 gives the rank discordance indices for the global list with the ranking lists from the US and Mexico. It is interesting to observe that the Mexico curve contains a single mode, whereas the US curve contains at least two modes, one at n=3, and the other around n = 10 (assuming the fluctuation around the second mode is random noises at this time).

In Figure 9, the discordance values for the five “Indo-European” countries – India, Pakistan, Germany, France, and Indonesia – are much larger than the corresponding values for the US and Mexico. One possible explanation may be that, in the Alexa data, the global ranking lists are weighted heavily toward US or North American consumers with the Alexa tool-bar, resulting in less discordance between the global and US/Mexico lists. Another observation is that the multimodal behavior in Figure 9 is even more pronounced. This probably indicates that the competition for the top few ranking positions is particularly sensitive to country-to-country variations, resulting in a greater discrepancy in “rank discordance” between the global list and those from this set of countries.

Figure 10 displays the discordance values for the global list with those of Japan and China. They also show a similar pattern as in the other cases. However, a much higher discordance is observed, which also occurs at a much earlier stage, signifying a stronger localization effect.
5. DISCUSSION OF THE FINDINGS
In this research, we examine the top-100 ranking lists of e-business Web sites from nine countries and compare them with the global top-100 list. The purpose is to determine whether we can observe any country-based differences among the consumers’ preferences in their choice of top e-business Web sites. The motivation behind this investigation emanates from the globalization and localization debate in international marketing.

The data used for this analysis are traffic-based Web site ranking lists obtained from Alexa.com. It is recognized that these Web site ranking lists contain noises because site visit statistics are collected from users who had voluntarily installed the Alexa tool bar on their browsers. They do not represent...
the entire online population of e-business consumers. As pointed out earlier, even
the concept of “site visits” on which the current data are based, may be subject to
several interpretations. The Alexa data are basically “reach” statistics (see
Section 3.1). They do not give us any indications on the “repeat-visit frequency”
or “stickiness” dimension. Another limitation we wish to acknowledge is that
Web site traffic data represents just part of (and not necessarily the whole of) the
e-commerce consumer’s online browsing behaviors. Nevertheless, we believe
that the data here may still be useful for the current analysis, provided we
interpret the findings with caution.

The metrics that we used in this research to measure the differences (or
similarities) between Web site ranking lists are based on an earlier work [Lo and
Sedheim, 2006], which was originally intended for investigating the reliability of
rank list providers. Three metrics were used to explore (a) membership
commonality, (b) rank correlation, and (c) rank discordance of Web site ranking
lists from different countries. The assumption is that differences (or similarities)
between these lists tell us something about consumers’ preferences in those
countries.

5.1. Discussion on Membership Commonality Results

A closer examination of the membership commonality matrix in Table 4
shows some interesting patterns. We modify the matrix in Table 4 and present it
here in Table 5, to include just the nine countries. We round up (or truncate) the
cell values to exaggerate their differences and similarities.

<table>
<thead>
<tr>
<th>Table 5: Membership Commonality Matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>us</td>
</tr>
<tr>
<td>mx</td>
</tr>
<tr>
<td>in</td>
</tr>
<tr>
<td>pk</td>
</tr>
<tr>
<td>ge</td>
</tr>
<tr>
<td>fr</td>
</tr>
<tr>
<td>id</td>
</tr>
<tr>
<td>jp</td>
</tr>
<tr>
<td>cn</td>
</tr>
</tbody>
</table>
The strong commonality (with a cell value of $\geq 0.40$) between the US and Mexico, and India and Pakistan is clearly evident. The next level, which has a cell value of around 0.30, includes four countries: India, Pakistan, Germany, and Indonesia. The common link here appears to be the “English” language factor. Although Germany and Indonesia have their own languages, English is still used freely throughout the countries, both in government and in education. Next is the 0.20 level that seems to involve France and certain pairings of Germany, Indonesia, and Japan. The most dissimilar lists are China and Japan. It would appear that both language and cultural factors contribute significantly to this pattern of dissimilarity.

These observations are reinforced in Figures 1 and 2, shown earlier, where the nine countries were initially separated into three groupings. And, then when combined with Figure 3, the middle group was further subdivided into two (or even three) subgroups, resulting in the following pattern:

- **North American Group** (US and Mexico)
- **Indo-European Group**
  - **Indo-Asia Subgroup** (India and Pakistan)
  - **Islamic Subgroup** (Indonesia and possibly Pakistan)
  - **Central Europe Subgroup** (Germany and France)
- **East Asia Group** (Japan and China)

The above classification appears to make sense if one takes into consideration language and cultural factors. These may even be able to map to, in a qualitative way, the national culture characteristics as defined by Hofstede [1996]. A more quantitative interpretation will be the subject of a future investigation; so, we do not propose to go beyond this qualitative statement at this stage. It is important to recognize that the above classification is uni-dimensional, but that the effect of culture and language factors on consumer choices and preferences are often multi-dimensional. Therefore, our classification is likely to be too simplistic in the current context, but it does serve as a useful first approximation.

In highlighting the differences, we should not overlook some of the common features among the curves in Figures 1 to 3. We observe that there is a greater degree of fluctuation or divergence among the different curves when the value of $n$ (the list size) is small, say $n = 1$ to about 3. This is not surprising because, when the list is small, if two countries differ (or agree) in their selection of the top 1, or top 2 (or even top 3) Web site(s), the $m$-value may fluctuate widely between 0 and 1. However, by the time the list size reaches about $n=5$, the membership commonality curves all seem to reach some sort of maximum value.
The actual maximum value may differ for different country-pairs. For example, the maximum m-value for the GL-US curve (Global list versus US list) peaks at about 0.8, and the maximum m-value for the DE-US curve peaks at about 0.6, whereas the maximum m-value for the IN-PK curve peaks at about 0.7. The most interesting observation is that, once the curves reach a maximum, they all seem to plateau off and sustain on the high m-value until around \( n=15 \) before they start to fall off. Remember that large m-value implies that two lists have higher commonality.

What this seems to indicate is that the Web site ranking list of two different countries are “most similar” in membership when the list size is around \( n=5 \) to \( n=15 \). Therefore, if for example, the consumers in two different countries were asked, “What do you think are the top-10 Web sites most frequently visited by your fellow country(wo)men?”, their answers are likely to be more similar compared with if they were asked the top-30, top-50 or top-100 Web sites. This commonality pattern seems to imply that the effect of globalization is most pronounced at the range of about \( n=5 \) to \( n=15 \). In another words, “famous” (or “good,” whatever that means) e-business Web sites (dominant transnational e-commerce giants) have established their “brand” position internationally and are recognized by most consumers regardless of from which country of origin the consumers come.

As \( n \) gets larger (beyond \( n=20 \)), the membership commonality curves in Figures 1 to 3 gradually decrease and appear to reach an asymptotic value when \( n \) becomes very large (in this case \( n=100 \)). However, the actual asymptotic values are different for different country-pairs. At \( n=100 \), the asymptotic values of m are in fact the membership commonality matrix of Table 4. Noting that the curves of Figures 1 to 3 have more or less completely leveled out at \( n=100 \), it is probably safe to say that the asymptotic m-value, denoted \( m_\infty \), is not likely to change further even as \( n \) increases. This means that, if we consider the top-200 Web site list, the value of \( m_\infty \) will probably remain the same. Therefore, we may say that this asymptotic m-value, \( m_\infty \), is no longer a function of \( n \) but is a characteristic of the two countries that we are comparing. This seems to imply that, as the ranking list gets larger (\( n \) increases beyond 20), consumers from different countries disagree more and more (smaller m-value) on who should be included in their “top” Web site list (whatever their concept of “top” means), until a stage is reached that the m curve reached an asymptotic value \( m_\infty \). The \( m_\infty \) value is therefore a measure of the “distance” between the localization characteristics of the two countries under consideration. This “distance” probably reflects the difference in culture, languages, and national conditions. Therefore, here we have a measure of the effect of localization.

5.2. Discussion on the Rank Correlation and Discordance Results

The second analysis involves the computation of Kendall’s rank correlation coefficients (\( \tau \)) for pairs of ranking lists from different countries. The results are presented in Figures 4 to 7, pairing each of the nine countries with the global list.
Remember that Kendall’s tau measures the extent to which two ordered-lists agree with each other in rank. Thus, it is a more discriminating measure than membership commonality. As a consequence, we expect that the “rank correlation” curves will show comparatively a lot more fluctuation than membership commonality curves, as their value is extremely sensitive to the “order” of the website in the ranking lists. This is clearly seen in Figures 4 to 7.

The next feature we observed in this set of figures is that, as the list size gets larger, the computed tau (τ) approaches different values for pairs of countries. There appears to be three or four sets of different asymptotic values for tau (τ): the US and MX are close to 0.6; IN, PK, and ID are close to 0.40 to 0.5; DE and FR are close to 0.4; whereas JP and CN are closer to 0.1 to 0.2. This lends further credence to the regional group classification proposed earlier.

Finally, we note that the tau-curves of seven of the countries (US, MX, IN, PK, ID, DE, and FR), except JP and CN, all cross over with the “critical-value” (tau*) curve when n becomes large (at around n=15). The tau curves for Japan and China show that, for these two countries, the consumers’ choices and preferences for the top-ranking Web sites are very different from those of the other countries. This may be attributed to the distinctive language and cultural differences for these two eastern Asian countries.

The third analysis involves the most discriminating metric, Δ, which measures the rank discordance for pairs of countries in the ranking lists. The results are presented in Figures 8 to 10. The most obvious observation is that the discordance from the global list is the smallest for the US and Mexico. This probably means that the “global” e-business ranking list is most likely strongly influenced by consumers in the US, as it is the most established country with respect to Internet e-businesses. Therefore, the US consumers’ choice and preferences are probably strongly reflected in determining the global ranking list. As time passes, some of the prominent US-based e-business sites will become better known in other countries, resulting in a convergence of these ranking lists. As indicated at the beginning of this paper, the effect of time on site ranking will be the subject of a separate study. Therefore, we shall restrict out comment to the “snapshot” picture here.

In the data presented here, the biggest deviation (rank discordance) was observed with the JP and CN group, confirming the comments we made earlier about their distinctive language and culture. The remaining countries (IN, PK, ID, DE, and FR) appear to have a discordance values that lie somewhere between the two extremes.

The most puzzling aspect of the curves in Figures 8 to 10 is the existence of more than one mode (2 or more maxima) for some of the countries. This is something that we hope to look for an explanation in the future to determine whether it is just a statistical “noise” or a genuine phenomenon.
6. CONCLUSIONS AND FUTURE RESEARCH
The purpose of this research is to gain a better understanding of the influence of the two forces, globalization and localization, in e-business cyberspace, by comparing the top-100 Web site ranking lists from different countries. Our analysis seems to confirm that both homogenization and fragmentation are occurring at the same time. In the complex e-business cyberspace, traditionalism and modernity are not necessarily mutually exclusive. Both forces continue and will continue to shape consumer decisions and behaviors. To some extent, e-commerce giants may be able to establish their “brand” dominance globally, but local culture and regional norms should not be ignored. Therefore, e-business professionals must take both forces into consideration as they plan their e-business strategies.

We recognize that the data sets used in this analysis give only an incomplete picture of reality. They also may contain noises that need to be carefully interpreted. Furthermore, the analyses performed in this paper are essentially qualitative and lack predictive power. It is therefore proposed that, in future research, one may include:

1. A comprehensive data set that includes more countries representing a wider variety of national cultural characteristics
2. A more comprehensive data set that includes a variety of e-business industries, as the nature of the e-business may have a strong impact on how the website is accepted by the consumers in different countries depending on local needs and demands
3. A series of data sets that are collected over time to enable the observation of temporal changes in site ranking. It is conceivable that some of the newer e-business sites may gain more popularity as they become better known in the developing countries. An example is the increasing acceptance of social networking sites among Asian countries.
4. A more quantitative model, with predictive power, to investigate the relationships of cultural and language factors and the ranking preferences of consumers. These models may include non-linear and/or multi-dimensional techniques.
REFERENCES


Alexa. 2007. www.Alexa.com Web site ranking data for different countries were retrieved from this site over a 12-month period from January to December 2007. The data set in this paper was collected in the week of September 7, 2007.


ABOUT THE AUTHORS

Bruce W.N. Lo is currently professor of information systems at the University of Wisconsin-Eau Claire. His first degree is from the University of London, his master’s degree is from the University of Newcastle, and his Ph.D. is from Monash University, Australia. He has taught and conducted research in Australia, Canada, China, the United Kingdom, and the United States. His research interests include software engineering, object technology, electronic commerce, IT education and management, and international business.

Faye J. Kao is assistant professor of marketing at Eastern Michigan University. Her previous experiences include account planning in an advertising agency, marketing management in a pharmaceutical company, and consulting in a marketing research agency and business consulting firm. Her current research interests are cross-cultural consumer behavior and international branding.