future internet

MDPI

Article

Environmental Factors Affecting Where People Geocache

Jennifer Golbeck [1,*] and Carman Neustaedter [2]

[1] College of Information Studies, University of Maryland, College Park, MD 20742, USA
[2] School of Interactive Arts and Technology, Simon Fraser University, Burnaby, BC V5A 1S6, Canada; carman@sfu.ca
* Correspondence: jgolbeck@umd.edu; Tel.: +1-301-405-7185

Academic Editors: Salvatore Carta and Ludovico Boratto
Received: 11 February 2016; Accepted: 5 April 2016; Published: 12 April 2016

Abstract: Outdoor leisure activities are important for public health as well as family cohesiveness, yet environmental factors may easily affect someone's ability to participate in such activities. We explored this with a focus on the social web-based treasure hunt game called Geocaching. We collected data on all US and Canadian geocaches from OpenCaching.com and conducted an online survey with twenty geocachers as a follow-up to our data analysis. Data analysis showed that geocaches were more often found in areas that were wealthier, better educated, younger, and more urban, and had higher population density and better weather. Survey results showed similar trends: Most people actively thought about where they would cache and tried to minimize risks, despite cache hiders thinking less about these concerns. These results further emphasize the importance of environmental factors when it comes to participation in outdoor activities and leads to Human–Computer Interaction design implications for location-based online social activities.

Keywords: geocaching; leisure activities; outdoors; physical activity

1. Introduction

Physical activity is considered an important public health priority to ensure overall health as well as life satisfaction. Leisure time through outdoor activities is also important for families and provides them with opportunities to bond, problem solve, and strengthen relationships [1]. Outdoor family recreation has also been found to help maintain and increase family cohesiveness [2]. Yet not everyone may have the same opportunities to participate in outdoor leisure activities because of various environmental factors, such as access to parks, poor weather at certain times in the year, and neighborhood crime.

Given this, we set out to answer the research question: What environmental factors affect where people participate in outdoor activities? In particular, we were interested in the GPS-based treasure hunt, Geocaching, which continues to grow in size and now has more than 6 million players worldwide.

Geocaching is a location-based treasure hunt game where players hide physical containers for other players to find using GPS coordinates posted online. Each container holds a logbook and optional tradable items. After players hunt for a geocache or "cache", they record their activity online by posting logs of "found", "did not find", *etc.* along with textual messages for other geocachers that describe their experience [3,4].

A number of studies have explored how and why people geocache. Motivations include making walks or hikes "purposeful", competing with friends, exploring nature, and participating in activities with family or friends [4]. Players hide geocaches for others as a way to bring people to unique places or places of meaning for the creator, or for the pure goal of creating a "challenge" for someone else [3]. These acts have created a community culture within the player base of geocaching where people report

on and help maintain the geocaches owned by others [3]. Studies have also shown that geocachers routinely hide and find geocaches in the areas they frequent, including near home, work, and in transit between the two [4]. We also know that people geocache while on vacation [3]. While valuable, what we do not know is if and how people are considering environmental factors when selecting where to geocache and how this might affect their participation.

The environmental impacts on physical activity have been studied from a variety of perspectives with the goal of understanding the relationship between an individual's preferences, behavior, and the environment [5]. There are a number of associations between personal attributes and physical activity. People who are older, live in rural areas, have less education, and have lower income have all been associated with less physical activity [6].

A review of literature in this area [7] assessed results from 19 different studies to determine correlations between activity and both perceived and objective physical environmental variables. They found that close access to places for walking, running, *etc.* (*i.e.*, parks and walking trails) were associated with more activity. Awareness of these opportunities as well as esthetically pleasing areas (friendly, scenic, attractive places) tended to increase activity. Weather seemed to have no impact, and most studies found no association with perceived neighborhood safety. This type of work inspired us to consider how environmental factors would impact participation in outdoor activities like geocaching.

Geocaching represents an interesting exploratory case for outdoor activities since it can be performed in nearly any outdoor location and people can be highly selective in terms of where they participate; the widespread placement of geocaches makes this possible.

To do this, we looked at two types of geocaching: hiding caches and searching for caches. We collected data on all the US and Canadian caches on OpenCaching.com available as of September 2013. To obtain environmental factors, we collected environmental information, including climate data, demographics, and crime rates on the Canadian census tract and US zip code level. Each cache was connected with the corresponding environmental data, and we analyzed the relationship between the environmental data and cache activity.

We found that environmental factors have almost no impact on cache *placement* with the exception of population density: There are more caches placed in less dense areas. However, almost every environmental factor we measured had a significant relationship with the number of cache logs (how often a cache was searched for). Caches with the most search activity tended to be in areas that were wealthier, better educated, younger, and more urban, and that had higher population density and better weather. We followed this with an online survey with 20 geocachers that further explores why they choose to hide and find caches in particular areas. Results showed that, while people were cautious in selecting areas to cache in (*i.e.*, that the areas were safe), they did not generally think about such concerns when hiding caches. Weather also played a factor in when and how people geocached.

Overall, these results show that it may be easier for people to participate in the outdoor activity of geocaching in certain locations. If such locations are not accessible to people, participation will be more limited. This further emphasizes the importance of environmental factors when considering when and how people may be able to participate in outdoor leisure activities and suggests that such activities may be need to be designed in somewhat different ways, depending on the location, to encourage broader participation.

2. Data Collection

2.1. Cache Data

We chose to examine caches in the US and Canada. Similar demographic information was available for both countries. Because the data was accessible via application program interface (API), we used cache data from OpenCaching, a free and open caching service. Our goal was to collect all the caches in the US and Canada from the OpenCaching dataset. To do this, we obtained the coordinates for 1217 US cities and 1619 Canadian postal code prefixes and requested the 5000 closest caches to

that location. After reducing the list to remove the large number of duplicates, we had a total of 23,318 caches in our dataset. The overlap in data from each query suggests that we did indeed collect all the caches in the US and Canada.

2.2. Demographic Data

We obtained demographic information on the zip code level for the US and on the census tract level for Canada. These correspond to similarly sized areas. The average area of a Canadian census tract in our data was 240 km^2 compared to 231 km^2 per zip code. Census tracts generally comprise between 2500 and 8000 people, and the average population per zip code in the US is roughly 7500.

For clarity, we will refer to zip codes and census tract areas simply as *zones*. There were a total of 4403 zones with caches in our data set.

For each zone, we obtained the following demographic information: climate data (average high and low temperatures in January and July, average precipitation), age, marital status (single, married, divorced, widowed), area, population, population density, education level, and income. We obtained most of this data from the 2011 Canadian Census and 2010 United States Census. Climate data for Canada was collected from the Canadian government's climate service and from the National Oceanic and Atmospheric Administration (NOAA) in the US.

Education level was broken down into the percentage of the population with some high school (SHS), a high school diploma (HSD), some college (SC), or a bachelor's degree or higher (B). We used these values to generate an "education" score for a zone by simply weighting each as follows:

$$1 \times \mathrm{SHS} + 2 \times \mathrm{HSD} + 3 \times \mathrm{SC} + 4 \times \mathrm{B} \quad (1)$$

We also collected crime risk data (from Homefair), although this was only available for the United States.

2.3. Data Integration

We wrote code to use the GPS coordinates of each cache to place it in the correct zone. Then, we associated each cache with the corresponding demographic data. In addition, we aggregated cache data by zone. For each zone, we had the total number of caches, the average number of finds and logs per cache, and average ratings of cache size, terrain, difficulty, and awesomeness.

3. Data Analysis

The number of caches per zone followed a power law distribution, with 2234 of the 4403 zones only having one cache. Similarly, the number of logs per cache in zones also followed a power law distribution, with 1959 zones averaging 1 find or fewer per cache.

We divided the zones into low- and high-cache zones and into low- and high-log zones. Low-cache zones had only one cache, while high-cache zones had more than one. Low-log zones averaged one log or fewer per cache, while high-log zones averaged more than one log per cache. This rough division allowed us to study what attributes corresponded to zones in each category.

We compared the values for different attributes between the high and low groups using a two-tailed Student's *t*-test in R. All significance values below are for $p < 0.05$. In addition, we confirmed results in both directions in the data. For example, if high-cache zones show a higher average for variable X than low-cache zones, then we confirmed that areas with a higher average for variable X indeed have a higher average number of caches. Below, we will only report the results in one direction for clarity, but the significance holds in both directions unless otherwise noted.

Note that all demographic features describe the *zones*, not necessarily the users who hid or found the caches.

The low- and high-cache zones shed insight into where people *hide* caches. We found that, among all the demographic data we collected, the only statistically significant difference between low- and

high-cache zones was the population density. Low-cache zones had significantly higher population density *vs.* high-cache zones (717 people/km^2 *vs.* 576 people/km^2). That is, caches are more often hidden in places with lower density.

However, looking at the number of logs per zone paints an entirely different picture. First, the population density result from the low- and high-cache zones was reversed. While high-cache zones have lower population density, high-log zones have higher population-density than low-log zones (778 people/km^2 *vs.* 485 people/km^2). This means people are logging more searches in higher-density areas.

Many other demographic attributes were significant as well:

Age: High-log zones have significantly younger populations (37.9 *vs.* 39.9 years old on average).

Climate: High-log zones were significantly warmer (averaging about 3 °F warmer than the measurements for low-log zones) and had less precipitation (an average of 41.6 mm less per year).

Area: High-log zones were significantly smaller than low-log zones (171 km^2 *vs.* 306 km^2), indicating more urbanness in high-log zones.

Education: The education score of the population was significantly higher in high-log zones (2.14 *vs.* 2.08).

Income: Median income was higher in high-log zones *vs.* low-long zones (about $6,170 higher).

Crime: Crime was significantly higher overall in the personal and property crime sub-categories in high-log zones *vs.* low-log zones. This is likely an indication that high-log zones tended to be more urban. There were no significant differences for overall crime and personal crime; higher crime areas had no significant difference in the number of logs compared to lower-crime areas.

The analysis of these zones reveals that, while lower population density areas tend to have more caches, higher population density areas tend to have more logs per cache.

Furthermore, while no demographic features significantly impacted the number of caches in a zone, nearly all of our demographic features showed significant differences based on the number of logs. Despite a very democratic placement of caches, those with the most activity tended to be in zones that were wealthier, better educated, younger, and more urban, and had better weather.

4. Online Survey Method

To better understand why people were hiding and searching for geocaches in particular areas, and if they thought about environmental factors when deciding where to geocache, we followed up our data analysis with an online survey. The survey was a mixture of open- and closed-ended questions that asked people when they geocached throughout the year, where they geocached, why, and if and how various environmental factors (e.g., weather, crime, perceptions of education levels, *etc.*) affected geocache hiding or searching.

We posted links to our survey on several geocaching forums across North America that varied in their environment (e.g., warmer weather *vs.* cooler, population density, *etc.*). We also posted a link to the survey on Twitter and Facebook. Twenty people responded with a range of geocaching experience. The median number of geocaches found by respondents was 1607 (range: 10–21,250). The median number of geocaches hidden by respondents was 16 (range: 0–400). Participants had a range of occupations (e.g., software engineer, stay-at-home mom, student, swim coach, *etc.*), income levels (from <$10,000 to a maximum of $160,000/year), education levels, and home locations (e.g., rural, suburb, *etc.*). Eight were from the US, ten from Canada, one was from the UK, and one from Germany.

We used thematic analysis and descriptive statistics to analyze our survey data. Despite the demographic variations amongst participants, we did not see any demographic-based trends in our survey results.

5. Online Survey Results

5.1. Searching for Geocaches

Nearly all respondents reported searching for geocaches throughout all months of the year; thus, season did not affect when they geocached. They also reported searching in all types of areas or neighborhoods, including rural, suburb, and downtown areas as well as parks, forests, and shopping areas. The only outlier was private property: only 45% of respondents reported searching for geocaches on private land areas (e.g., homes and yards). Private property was considered a questionable location to geocache because people did not know if the landowners granted proper permission for the activity.

Weather affected when 95% of respondents would geocache. Respondents said they would avoid geocaching during rain, snow, very cold weather, and very hot weather. This sometimes related to environmental factors affected by the weather such as increased amounts of insects or increased terrain difficulties due to wet conditions.

"I like going after bush caches in the late fall—no bugs, snow hasn't arrived yet, brush is thinned out. I like winter caching for the bug-free part, fewer muggles, geotrails, brush gone, as long as they're winter-friendly."—P12

Perceived crime rate of an area affected where 75% of respondents geocached. These individuals were very careful to think about the potential risk associated with caching in areas that they thought had a high crime rate. Sometimes they would restrict caching to daylight hours as a result or avoid an area altogether.

"I used to cache ALMOST anywhere, until a friend mentioned to me that the places I'd been taking my kids are really really bad areas with high crime rates. I stopped immediately, and starting paying more attention to the areas of town that friends told me about. I am much more aware now."—P2

In contrast, the type of people found in an area (e.g., areas with perceived lower or higher income levels or education) did not affect geocache searching, but some respondents did report not going to areas that contained people suffering from substance abuse issues (e.g., drugs, alcohol).

5.2. Hiding Geocaches

Hiding geocaches was a less frequent activity; thus, people performed it at varying times during the year. No months were more dominant than others for hiding caches. The most popular place for hiding a geocache was near someone's home (60% reported doing this), followed by someone's route between home/work (30%), near friends/family (20%), vacation or travel destinations (20%), and other locations not frequented (15%). None of our respondents reported placing caches on private property.

Respondents thought about three main factors when hiding geocaches: how long it took to get to the location in order to maintain the cache, the number of "muggles" (non-cachers) present, and the density of caches in the area (trying not to "overpopulate" an area). They tried to reduce challenges in each of these cases. Weather affected when 55% of respondents hid geocaches. Again, they tried to avoid hiding during particularly "bad" weather conditions such as when snow was present. However, they did not comment on trying to reduce weather issues for those searching for their geocaches. Only 30% of respondents thought about perceived crime in the areas they hid geocaches and would choose to hide them in other areas. Because people tend to hide geocaches in areas that are familiar to them, it is likely that perceived issues of safety or crime in an area are minimal because of the location's familiarity rather than the true crime rate.

6. Discussion and Conclusions

Our results show that participation in geocaching is most likely to occur in certain areas, including places that are wealthier, better educated, younger, and more urban, and that had higher population density and better weather. When asked, respondents clearly think about the types of locations that they geocache in by trying to reduce areas containing high crime. They also chose to cache in "good weather" (e.g., no rain, snow, extreme hot or cold). This favors locations with a moderate year-round

climate. Overall, these results show that if people happen to live in an area that fits the desirable environmental factors, it is likely that they will have more opportunities to participate in geocaching and find it more enjoyable. Poorer, less-educated, rural, cold and rainy areas still have access to caches, but they are not as frequently sought out.

This work complements other research on geocachers that capture their personal attributes and attitudes [8]. Our results also echo those of [9], which found that younger, more educated people who live in more dense urban areas tend to be geocachers.

This further emphasizes the importance of environmental factors when considering when and how people may be able to participate in outdoor leisure activities. To encourage participation in locations that are "less popular" geocaching destinations, people should think about the way the activity is presented and, perhaps, design it more specifically for the region to encourage broader participation. For example, in areas with low average incomes, people could consider ways of supporting geocaching where an expensive GPS device or smartphone is not required. In areas containing poor weather, people could think about, for example, supporting caches that still require physical activity but may be partially indoors. More generally speaking, people could consider similar ideas when creating or promoting physical and leisure activities by designing programs that take the environmental factors that are likely to affect participation into account.

Author Contributions: J.G. collected the data; J.G. and C.N. analyzed the data; J.G. and C.N. wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Agate, J.; Zabriskie, R.; Agate, S.; Poff, R. Family leisure satisfaction and satisfaction with family life. *J. Leis. Res.* **2009**, *41*, 205–223.
2. West, P.; Merriam, L.C. Outdoor recreation and family cohesiveness: A research approach. *J. Leis. Res.* **2009**, *41*, 351–359.
3. Neustaedter, C.; Tang, A.; Judge, T. The Role of Community and Groupware in Geocache Creation and Maintenance. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Atlanta, GA, USA, 10–15 April 2010.
4. O'Hara, K. Understanding Geocaching Practices and Motivations. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Florence, Italy, 5–10 April 2008.
5. Bandura, A. *Social Foundations of Thought and Action: A Social Cognitive Theory*; Cliffs, N.J.E., Ed.; Prentice-Hall: Upper Saddle River, NJ, USA, 1986.
6. Wilcox, S.; Bopp, M.; Oberrecht, L.; Kammermann, S.; McElmurray, C. Psychosocial and perceived environmental correlates of physical activity in rural and older African American and white women. *J. Gerontol. Ser. B Psychol. Sci. Soc. Sci.* **2003**, *58*, 329–337. [CrossRef]
7. Humpel, N.; Owen, N.; Leslie, E. Environmental factors associated with adults' participation in physical activity: A review. *Am. J. Prev. Med.* **2002**, *22*, 188–199. [CrossRef]
8. Telaar, D.; Krüger, A.; Schöning, J. A large-scale quantitative survey of the german geocaching community in 2007. *Adv. Hum. Comput. Interact.* **2014**, *2014*. [CrossRef]
9. Cord, A.F.; Roeßiger, F.; Schwarz, N. Geocaching data as an indicator for recreational ecosystem services in urban areas: Exploring spatial gradients, preferences and motivations. *Landsc. Urban Plan.* **2015**, *144*, 151–162.