Cognitive Assessment of Geospatial Survey Data

David J. Roe, MA[1], Douglas Currivan, PhD[1], James B. Holt, MPA, PhD[2], George Ghneim, DVM, MPVM, PhD[1]

[1]RTI International, Research Triangle Park, NC; [2]Centers for Disease Control and Prevention (CDC), Atlanta, GA

Abstract

Mapping survey data has become an increasingly important means of displaying survey information and highlighting differences across geographic areas. In public health, geospatial data are used by policy-makers, researchers, and other health professionals to identify and address issues of public health concern.

Traditionally, geospatial data are presented in choropleth maps, where defined geographic units (i.e., political boundaries) are filled with a uniform color or pattern. Such maps are appropriate for data that have been scaled or normalized. For example, high school graduation rates displayed by county are typically represented using choropleth maps. Alternatively, geospatial data can be displayed using isopleth maps, in which the data are not aggregated to pre-defined geographic units, but rather are typically "smoothed" across adjacent geographic boundaries. There is little empirical research, however, on the differential effectiveness of choropleth versus isopleth maps. In particular, how do these two different mapping techniques affect the user's ability to extract information from the map?

Using maps of health data collected from the Behavioral Risk Factor Surveillance System (BRFSS), a widely used source of public health information, we present the results from a series of focus groups and cognitive testing sessions which were used to evaluate the way in which public health professionals and epidemiologists commonly use mapped data. Among the issues explored are usability, general preferences, ease of pattern recognition, and rate retrieval for both choropleth and isopleth maps. A clear majority of participants preferred the isopleth format, however, their assessment varied somewhat by the data retrieval task. The color scheme used also had an impact, regardless of the type of map used. Implications for the assessment of geospatial presentation of survey data are also discussed.

Background

The Behavioral Risk Factor Surveillance System (BRFSS), which was established in 1984 and reached national state-level data collection in 1994, was developed "to collect prevalence data on risk behaviors and preventive health practices that affect health status" (Centers for Disease Control and Prevention [CDC], 2005). BRFSS survey data are collected annually, and results are used for health promotion and disease prevention program planning, development, and evaluation. Certain results are presented in the form of GIS Maps illustrating health risks at national, state, and local levels. The BRFSS survey consists of a core interview (i.e., a fixed core, rotating core, and emerging core questions that reflect "late breaking" issues), optional modules, and state-added questions. RTI International conducts a variety of BRFSS activities for the CDC, including survey item appraisal and cognitive testing of items.

BRFSS
Behavioral Risk Factor Surveillance System

Methods

Cognitive assessment for the BRFSS mapping study included (1) a focus group in Research Triangle Park, North Carolina, and (2) two rounds of in-person cognitive interviews, one in Research Triangle Park and one in Atlanta, Georgia. Participants in all three groups were presented with a series of maps accompanied by multiple choice questions focused on rate retrieval, pattern recognition, and ease of understanding. Additional questions asked about participants' preference for different map formats and which format they found easier to read.

Two sets of maps were used for the cognitive interviews, six-color maps and four-color maps. Each set of maps was divided between choropleth and isopleth maps. ***Choropleth*** maps present data as geographic areas shaded with intensity proportional to the data values associated with those areas. ***Isopleth*** maps display data smoothed across the midpoints for each area, regardless of geographic boundaries. Because county boundaries shown on these maps affected users' ability to retrieve data, the thickness of the boundary lines was also assessed in cognitive interviews.

Methods (cont'd)

The focus group protocol was designed to elicit reactions from public health professionals and researchers on the different BRFSS maps and the questions asked about the maps. We expected the focus group discussion to contribute to the mapping experiment by (1) examining assumptions about the mapped data and map questions that participants would be asked to use and (2) evaluating assumptions about the way participants understand key terms or concepts used with mapped data.

The cognitive interview protocol used retrospective verbal probing to ask participants about using each map and answering the questions. Probing was conducted after all map questions were answered, to ensure that discussions between participants and interviewers did not influence participants' answers to subsequent questions. These interviews elicited feedback about each map focusing on:

- Question comprehension
- Retrieval from memory
- Decision processes
- Response processes

What Do You Think?

Take a look at the top two maps on the right. Can you find the correct answers to the questions below? As you answer each question, ask yourself, "Would this question be easier or harder to answer on the other map?"

US General Health Choropleth (Top Left)

What is the predominant rate of adults reporting having fair or poor health in Nevada?

a. Less than 13.6%
b. 13.6%–18.4%
c. 18.5%–24.9%
d. Cannot assess

US General Health Isopleth (Top Right)

Which 3 states have an area where there is a clear gradient (a stepwise increase) involving all 4 of the health rates?

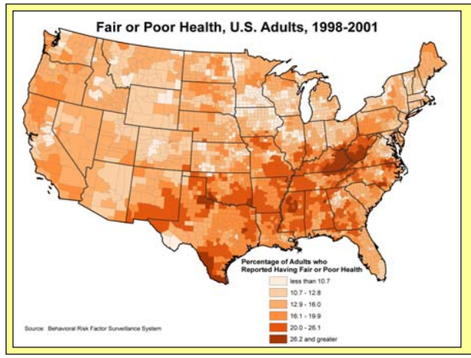
a. Ohio, North Carolina, and Virginia
b. Ohio, North Carolina, and Kentucky
c. Ohio, North Carolina, and Georgia
d. Ohio, Kentucky, and Virginia

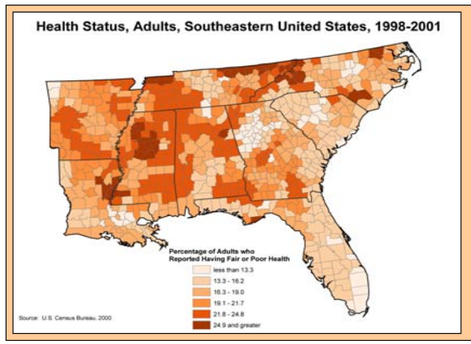
Conclusions

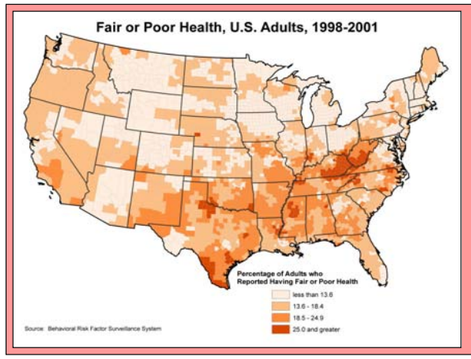
Incorporating the results of the focus group and the two rounds of cognitive testing provided several important conclusions.

- A clear majority of participants preferred the isopleth format over the choropleth format and found the isopleth format easier to read. Results and comments indicate that this was particularly true when looking at the U.S. maps, but the majority chose the isopleth format in reference to both the U.S. and Southeast maps.
- Overall, retrieving data from U.S. maps was more difficult than retrieving data from Southeast regional maps when the maps were in choropleth format, six-color format, or both. Retrieving data from choropleth U.S. maps was especially difficult in the eastern two-thirds of the country.
- Although participants generally preferred the isopleth maps, their assessment varied somewhat by the data retrieval task. For pattern recognition, participants clearly preferred the isopleth maps across all formats. For rate retrieval, participants found the choropleth maps as useful as or more useful than the isopleth maps in some situations, such as discerning county rates from a Southeast regional map.
- Participants had greater difficulty distinguishing color categories with the six-color maps compared to the four-color maps. This was particularly true for the middle categories as opposed to the lowest and highest categories.
- On choropleth maps, discerning county boundaries was more difficult for respondents with the darker boundary outlines compared with the lighter boundaries. This was especially true when retrieving data from smaller states, states with relatively many counties, along state borders, or along coastal areas.

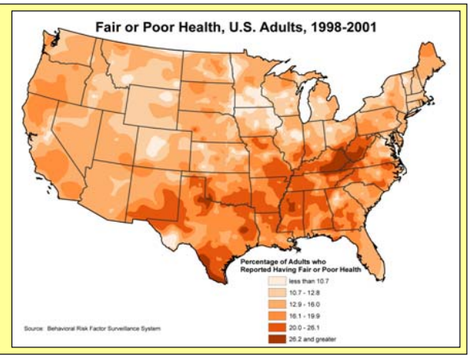
Choropleth

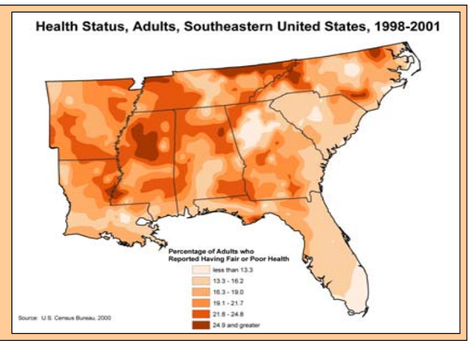


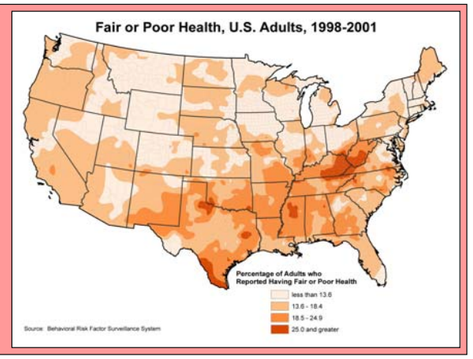




Isopleth

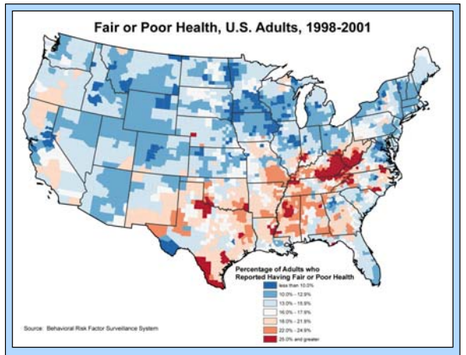


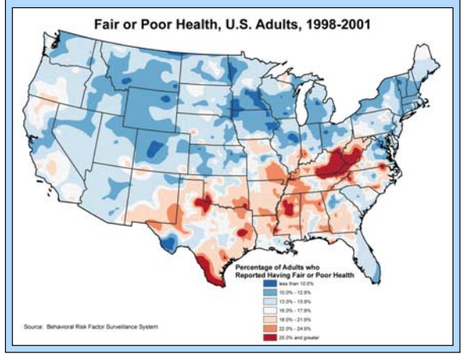




Next Steps

The results of cognitive interviewing demonstrated that respondents had considerable difficulty differentiating color categories based on the orange color ramp, especially when using the six-color maps. These results supported further investigation and testing of other color ramps that provide a greater contrast between classes in a given ramp. Also, after seeing the difficulties respondents had with the dark boundary choropleth maps during the focus group and cognitive tests, we concluded both choropleth and isopleth maps should include light county boundaries. The two new maps below show the same general health data as the first two maps to the left, with a new divergent color ramp, which goes from red to blue. The new map format, designed to improve user accuracy in retrieving data from choropleth and isopleth maps, will be tested through an internet survey of approximately 400 public health researchers and professionals. Data from the internet survey will facilitate analysis of key issues such as users' accuracy of data retrieval from choropleth versus isopleth maps, differences in data retrieval from national versus regional maps and users' preferences for choropleth versus isopleth maps. These results will allow the CDC to determine the best format and features for providing online maps for BRFSS data users.





Contact the Authors:

David Roe: droe@rti.org
Douglas Currivan: dcurrivan@rti.org
James B. Holt: jgh4@cdc.gov
George Ghneim: gsghneim@rti.org

Presented at the 61st Annual American Association for Public Opinion Research (AAPOR) Conference, Montreal, Quebec, May 18th–21st, 2006