NEURO-IMAGING SUPPORT FOR THE USE OF AUDIO TO REPRESENT GEOSPATIAL LOCATION IN CARTOGRAPHIC DESIGN

by

MEGEN ELIZABETH BRITTELL

A DISSERTATION

Presented to the Department of Geography
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

December 2018

DISSERTATION APPROVAL PAGE

Student: Megen Elizabeth Brittell

Title: Neuro-imaging Support for the Use of Audio to Represent Geospatial Location in Cartographic Design

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Geography by:

Amy Lobben	Chair
Patrick Bartlein	Core Member
Hedda Schmidtke	Core Member
Michal Young	Institutional Representative

and

Janet Woodruff-Borden	Dean of the Graduate School

Original approval signatures are on file with the University of Oregon Graduate School.

Degree awarded December 2018

DISSERTATION ABSTRACT

Megen Elizabeth Brittell

Doctor of Philosophy

Department of Geography

December 2018

Title: Neuro-imaging Support for the Use of Audio to Represent Geospatial Location in Cartographic Design

Audio has the capacity to display geospatial data. As auditory display design grapples with the challenge of aligning the spatial dimensions of the data with the dimensions of the display, this dissertation investigates the role of time in auditory geographic maps. Three auditory map types translate geospatial data into collections of musical notes, and arrangement of those notes in time vary across three map types: *sequential*, *augmented-sequential*, and *concurrent*. Behavioral and neuroimaging methods assess the auditory symbology. A behavioral task establishes geographic context, and neuroimaging provides a quantitative measure of brain responses to the behavioral task under *recall* and *active listening* response conditions.

In both behavioral and neuroimaging data, two paired contrasts measure differences between the sequential and augmented-sequential map types, and between the augmented-sequential and concurrent map types. Behavioral data reveal differences in both response time and accuracy. Response times for the augmented-sequential map type are substantially longer in both contrasts under the active response condition. Accuracy is lower for concurrent maps than for

augmented-sequential maps; response condition influences direction of differences in accuracy between the sequential and augmented-sequential map types. Neuroimaging data from functional magnetic resonance imaging (fMRI) show significant differences in blood-oxygenation level dependent (BOLD) response during map listening. The BOLD response is significantly stronger in the left auditory cortex and planum temporale for the concurrent map type in contrast to the augmented-sequential map type. And the response in the right auditory cortex and bilaterally in the visual cortex is significantly stronger for augmented-sequential maps in contrast to sequential maps. Results from this research provide empirical evidence to inform choices in the design of auditory cartographic displays, enriching the diversity of geographic map artifacts.

Four supplemental files and two data sets are available online. Three audio files demonstrate the three map types: sequential (Supplementary Files, Audio 1), augmented-sequential (Supplementary Files, Audio 2), and concurrent (Supplementary Files, Audio 3). Associated data are available through OpenNeuro (https://openneuro.org/datasets/ds001415).

CURRICULUM VITAE

NAME OF AUTHOR: Megen Elizabeth Brittell

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, Oregon
University of Maryland, Baltimore County, Baltimore, Maryland
Johns Hopkins, Engineering Program for Professionals, Elkridge, Maryland
Washington State University, Pullman, Washintgon
Willamette University, Salem, Oregon

DEGREES AWARDED:

Master of Science, Computer and Information Science, 2011, University of Oregon
Bachelor of Arts, Mathematics and Russian Studies, 2002, Willamette University

AREAS OF SPECIAL INTEREST:

Spatial Cognition
Cartography
Human Computer Interaction

GRANTS, AWARDS AND HONORS:

Doctoral Dissertation Research Improvement Grant (#1634068), National Science Foundation, 2016

Doug Foster Community-Building Award, Department of Geography, University of Oregon, June 2016.

American Association of Geographers, Environmental Perception and Behavioral Geography Student Travel Award, April 2016.

International Cartographic Association, US National Committee student travel award, International Cartographic Conference, August 2015.

University of Oregon Women in Graduate Science travel award, International Cartographic Conference, August 2015.

Conference on Spatial Information Theory, Doctoral Colloquium Student Award (University of Saskatchewan), October 2015.

North American Cartographic Information Society, Student Travel Award, October 2014.

ACM Special Interest Group on Spatial Information (SIGSPATIAL), Runners-up Best Demonstration, November 2013.

J. Donald Hubbard Family Scholarship, University of Oregon Department of Computer and Information Science, 2011

Magna Cum Laude, Willamette University, 2002

PUBLICATIONS:

Brittell, M., Lobben, A., and Lawrence, M. (In Press 2018). Evaluation of Tactile Map Symbols across Three Production Technologies. *Journal of Visual Impairment & Blindness*.

M. Brittell. (2018, June). Seeking a reference frame for cartographic sonification. International Conference on Auditory Display, Houghton, Michigan. Available: http://hdl.handle.net/1853/60082.

Brittell, M., Grummon, C., Lobben, A., Omri, M., and Perdue, N. (2017, May). Accessibility in Pedestrian Routing. In M. Peterson (Ed.), *Advances in Cartography and GIScience.* ICACI 2017 (pp. 479–492). Lecture Notes in Geoinformation and Cartography. Springer International Publishing. Cham, Switzerland. doi: http://dx.doi.org/10.1007/978-3-319-57336-6_33.

Lobben, A., Brittell, M., and Perdue, N. (2015, August) Inclusive Cartographic Design: Overcoming Ocular-Centric Cartographies. In C. Robbi Sluter, B. C. Madureira Cruz, & M. P. Leal de Menezes (Eds.), *Cartography - maps connecting the world*. ICACI 2015. (Chap. 7, pp. 89–98). Springer International Publishing. doi: http://dx.doi.org/10.1007/978-3-319-17738-0_7.

Brittell, M., Young, M., and Lobben, A. (2013, November). The mGIS: A minimal geographic information system accessible to users who are blind. Demonstration Paper, ACM Special Interest Group on Spatial Information (SIGSPATIAL), Orlando, FL. doi:http://dx.doi.org/10.1145/2525314.2525329.

Brittell, M. (2011, May). Line following: a path to spatial thinking skills. In *Extended Abstracts on Human Factors in Computing Systems, CHI EA '11*, Vancouver, British Columbia. doi: http://dx.doi.org/10.1145/1979742.1979840.

ACKNOWLEDGEMENTS

I would like to thank my dissertation advisory committee, Dr. Amy Lobben, Dr. Patrick Bartlein, Dr. Hedda Schmidtke, and Dr. Michal Young, for their guidance and support of my research. In particular, I would like to thank Dr. Amy Lobben and Dr. Michal Young for their support throughout my graduate studies, at both the Masters and Doctoral levels. Thanks also to the members of the former Spatial and Map Cognition Research Lab (SMCRL) and current Spatial Cognition, Computation, and Complexity Lab (S3C) for their technical, social, and moral support.

I would like to thank the staff of the Robert and Beverly Lewis Center for Neuroimaging (LCNI) for their assistance with planning and conducting data collection, and special thanks to Jolinda Smith for her assistance selecting and evaluating a suitable function scan sequence that accommodated the presentation of the auditory stimuli.

The work was improved by formal and informal feedback from study participants, and informal feedback from attendees at conference presentations and community members at local coffee shops.

This material is based upon work supported by the National Science Foundation under Doctoral Dissertation Research Improvement Grant No. #1634068. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

This research was supported in part by the University of Oregon Lewis Family Endowment.

Harvard-Oxford Cortical Structural Atlas, which is available in FSL, was provided by the Center for Morphometric Analysis (CMA) at Massachusetts General Hospital. We are very grateful for the training data for FIRST, particularly to David Kennedy at the CMA, and also to: Christian Haselgrove, CMA, Harvard; Bruce Fischl, Martinos Center for Biomedical Imaging, MGH; Janis Breeze and Jean Frazier, Child and Adolescent Neuropsychiatric Research Program, Cambridge Health Alliance; Larry Seidman and Jill Goldstein, Department of Psychiatry of Harvard Medical School; Barry Kosofsky, Weill Cornell Medical Center.

The Jülich Histological Atlas, also available in FSL, was developed by Zilles and Amunts at the Research Center Jülich and provided by Simon Eickhoff.

This work benefited from access to the University of Oregon high performance computer, Talapas.

Dissertation manuscript formatting in LaTeX supported by the template `uothesis.cls` based on work by Mark VandeWettering with modifications from Chet Haase, Christian Frank, Lars Hansen, J. Paul Walser, Bei Li, Bart Massey, et al., as well as collected revisions for University of Oregon graduate students via Rose Maier Hartman. Table formatting supported by the `xtable` package in R (https://CRAN.R-project.org/package=xtable; v1.8-2).

To Mom, Dad, and Maisy

TABLE OF CONTENTS

SUPPLEMENTAL FILES

AUDIO 1: SEQUENTIAL MAP (sequential.wav)

AUDIO 2: AUGMENTED-SEQUENTIAL MAP (augmented.wav)

AUDIO 3: CONCURRENT MAP (concurrent.wav)

LIST OF FIGURES

Figure Page

LIST OF TABLES

Table Page

CHAPTER I

INTRODUCTION

Auditory display offers a communication medium for multi-dimensional data, with the potential to extend and augment established graphic cartographic techniques. Adopting auditory display into the cartographer's toolbox may provide several benefits. Human hearing, which detects and processes auditory input, possesses multiple pattern detection capabilities (e.g., Wojtczak, Mehta, & Oxenham, 2017). Such capabilities support detection of patterns in data that may not be apparent in speech-based (e.g., Zhao, Plaisant, Schneiderman, & Duraiswami, 2004) or visual (e.g., Diaz-Merced et al., 2011) displays. Creating auditory displays also supports inclusive design principles (see Lobben, Brittell, & Perdue, 2015); as a non-visual modality, audio enhances access to geospatial data by people who are blind or low vision. Finally, auditory displays provide an opportunity to investigate spatial thinking and cartographic design with less susceptibility to visual bias. Auditory geographic map design challenges vision-based assumptions about inherent properties of geospatial data, countering ocular-centric momentum that cartographic theory and guidelines traditionally couch in visual terms (MacEachren, 1995; Robinson, Morrison, Muehrcke, Kimerling, & Guptill, 1995; Wood, 1968).

Several challenges, however, accompany these potential benefits and auditory geographic map design cannot overcome them without a better understanding of how audio can communicate the spatial dimensions of geospatial data. Existing tools to implement auditory displays target a user group composed of sound engineers and digital music composers who possess a specialized skill set

(Goudarzi, 2016; Kramer et al., 1997). But such skills are neither common nor widely taught among geographers and cartographers. Identification of necessary functionality to support auditory cartography is necessary to bridge the gap between domain expertise in geospatial data and specialist skills in sound design. Design guidelines to underpin auditory-map creation lag behind those of their visual counterparts. Published guidelines describe design of auditory computer interfaces and sonification of general scientific data (e.g., Brewster, Wright, & Edwards, 1995; L. Brown, Brewster, Ramloll, Burton, & Riedel, 2003; Dubus & Bresin, 2013), but do not address unique needs of geospatial data (Brittell, 2018). While several approaches to auditory map design have been proposed (see Chapter II: Background), consensus has not been reached on a definitive, usable solution. Geographic concepts behind visual- and tactile-map-design choices similarly apply to the auditory domain (e.g., distinction between topography, topology, and time in transportation maps Tatham, 1995), but techniques of modality-specific representation do not. Not only are there fundamental differences between the human perceptual systems, but using a visual model to inform auditory design detrimentally biases design choices (Frauenberger, Putz, Hoeldrich, & Stockman, 2005). However, advances in interdisciplinary research in sonification (Hermann, Hunt, & Neuhoff, 2011; Kramer et al., 1997) and methods for the study of spatial cognition (e.g., neuroimaging, Lobben, Lawrence, & Olson, 2009) create an opportune research setting to investigate cartographic applications of auditory display.

To support and facilitate further adoption of auditory display into the cartographer's toolbox, this dissertation describes an empirical study that investigates the representation of geospatial data in audio by exploring the role of time in auditory thematic maps. Seeking to understand how temporal arrangement influences uptake of information from an auditory map display, this research addresses two research questions:

RQ1: How does the temporal aspect of auditory map symbolization influence effectiveness in communicating general spatial patterns in the data?

RQ2: How does neural activation in response to serial audio symbolization of a geographic map contrast with that of simultaneous audio symbolization?

This research takes an approach that differs from previous work in three major ways. First, the proposed auditory map symbology challenges a simplistic use of time. Existing approaches to the display of spatial data in audio adopt time-based symbology: Meijer (1992) "distribute an image in time" encoding one spatial dimension in time, and Zhao, Plaisant, Shneiderman, and Lazar (2008) "temporalize" geospatial data to encode two spatial dimensions in time. While time is a strong dimension of audio, the argument presented here is that, as a multidimensional display modality, an auditory display can depict geospatial data without reducing it to a linear sequence. Second, the production of usable auditory maps relies on a design process that starts with data rather than translating a graphic display. A direct connection between an auditory map and its underlying data promotes auditory displays to first-class citizens among cartographic products. And third, the evaluation employs neuroimaging methodology that avoids (some) visual bias in design and evaluation of auditory maps. While experience using graphic maps and training in graphic map design encourages comparison with visual displays and taints design choices, brain activation reveals differences in neural responses to auditory maps that listeners need not consciously recognize or be able to articulate.

This manuscript provides a review of the literature in Chapter II: Background. Interdisciplinary contributions spanning geography, psychology, neuroscience, and computer science provide a foundation for the work. The methods detailed in Chapter III: Methodology describe the design of three auditory map types and their evaluation with a combination of behavioral and neuroimaging techniques. A behavioral measure provides insight into effectiveness of the auditory map display, while neuroimaging assesses differences in neural activity in response to the temporal arrangement of audio streams and the information that they encode. Chapter IV: Results presents results of the empirical study and Chapter V: Discussion interprets those results. Beyond confirmation that behavioral differences exist or that one specific design was different from another, the results provide evidence of differing patterns and intensities of brain responses across map designs. Chapter VI: Conclusions describes implications of this work. The manuscript ends with a collection of appendices that provide details of study implementation.

CHAPTER II

BACKGROUND

Investigation of how the temporal aspect of audio-based map designs influence cognitive processes is necessarily an interdisciplinary activity (using "interdisciplinary" as defined by Baerwald, 2010). This chapter traverses traditional disciplinary boundaries and draws from literature across multiple academic disciplines. Geography is a natural starting point for this discussion, contributing literature that provides theoretical frameworks to organize geospatial data ranging from digital data formats (e.g., Peuquet, 1984, 2001), to spatial cognition (e.g., P. J. Gersmehl & C. A. Gersmehl, 2007; Golledge, 2002). Further, theories from cartography contribute techniques for the representation and display of geospatial data to support cartographic communication. Computer science offers conceptual models and software tools for data storage, processing, and transmission by machines. And research in human computer interaction and human factors informs the design of computer-based displays to help humans extract information and produce knowledge from digital data (e.g., Schneiderman & Plaisant, 2010; Zhao, B. K. Smith, Norman, Plaisant, & Schneiderman, 2004). Psychology literature informs abstract and applied models of the human side of that interaction with a robust understanding of the human perceptual systems (Goldstein, 2014). The subfield of psychoacoustics (Bregman, 1990; Cook, 1999) offers evidence of the capabilities and limitations of the human sensory systems, and helps explain connections between stimulus characteristics and observable behavioral responses. Neuroscience, then, explores mechanisms of human cognition and behavior from the perspective of brain structure and neural connectivity. Relatively new methodologies from neuroscience, such as functional magnetic resonance imaging (fMRI, see Bandettini, 2012; Ogawa, Lee, Kay, & Tank, 1990), offer quantifiable metrics of human responses that reduce (some forms of) bias, which arise in evaluation of unfamiliar geographic map types. With the preceding general overview of contributions by the various selected disciplines, the literature review continues without explicit distinctions between the traditional disciplines.

The combined literature provides both theoretical support and empirical evidence across four general themes: geospatial data, representation, sonification, and the current state of audio cartography. These themes form a foundation for investigating auditory geographic maps. Geospatial data and their representation in geographic maps establish the scope of the application domain. The focus on geospatial data highlights the spatial reference frame in which non-spatial data is positioned and which makes maps a unique case of multi-dimensional data display. Digital formats, mental representations, and geographic map artifacts constitute distinct representations of geospatial data. The review describes each representation in turn before examining relationships between them. Narrowing the scope of the inquiry to thematic geographic maps, the narrative continues with a more extensive examination of sonification and uses of audio in cartography. Within the limited existing research on audio-only geographic maps, examination reveals a tendency to reduce spatial data to a linear stream while the necessity and implications of that design choice remain unclear. The chapter concludes by summarizing the current state of auditory cartography. This summary identifies a gap in the literature around the use of time in audio cartography and outlines the contribution that this research makes to help address it.

Geospatial Data

Geospatial data stem from observations of the physical and cultural world that are characterized by their association with geographic space, which spans a range of spatial extents from the immediate surroundings of an individual to the entirety of the earth's surface. Observations associated with a location in space make geospatial data necessarily multidimensional, and distinct from other multidimensional data sets (Brittell, 2018; Skupin & Fabrikant, 2003). Spatial dimensions create a reference frame in which to arrange non-spatial data values. To set the stage and scope for investigating its auditory representation, this section introduces geospatial data by describing one approach to organizing data in conceptual categories, outlining dimensionality, and putting forward a definition of spatial pattern.

Characteristics of geospatial data belong to three conceptual categories: space, time, and attribute. These categories occur under many labels. Berry (1964) adapts a structure from anthropology to organize "geographic facts" by *place*, *time*, and *characteristic*. MacEachren, Wachowicz, Edsall, Haug, and Masters (1999) refer to *where*, *when*, and *what* components of data and "space-time-attribute patterns." Lobben (2003) uses three criteria to categorize data: *space*, *time*, and *variable*. D. Guo, Chen, MacEachren, and Liao (2006) describe a "spatio-temporal and multivariate data cube" with labels *space*, *time*, and *multiple variables*. And, in human geography, Keller, Buck, Zare, and Popescu (2014) describe a "human geographic data cube" with "physical (i.e., terrain) and human (e.g., income, political, and cultural) variables" to predict future events, implicitly occurring in or over time. Terminology varies, but each of these examples propose similar conceptual categorization of data. And, notably, all include a category for space and refer to time; these two conceptual categories within geospatial data play fundamental roles in this exploration of auditory maps.

The three categories that make up the spatial data cube comprise multiple data dimensions. Each dimension provides a basis for measurement ("level of measurement", MacEachren, 1994; "measurement scale", Dent, 1999; "level of organization", Bertin, 1967/2011). For example, and of particular interest in geography, ordinal or ratio axes measure space, which is inherently relative (P. J. Gersmehl & C. A. Gersmehl, 2007; Montello & Raubal, 2012; Richter & Winter, 2014; Skupin & Fabrikant, 2003). Although relatively rare, space may be a single dimension. For example, the political boundaries of states may be used as the critical criteria for segmenting a data set, and a single nominal variable (state name) may be used to organize data (e.g., D. Guo et al., 2006). More commonly, the physical world embodies our conception of the space category, and data values on each of two or three dimensions describe locations in physical space (Peuquet, 1984). The two (surface of the earth) or three (surface of the earth and elevation) dimensions of physical space are orthogonal; geographic polar coordinates or projected Cartesian coordinates uniquely identify location in physical space. Higher dimensions of space are plausible (Peuquet, 1984), but largely isolated to theoretical discussion. The time axis from the data cube structures our experience with the world, but is unique in that it does not have a physical presence (Vasiliev, 1996). In contrast to space, time proceeds in a single direction, and the nature of the time dimension may preclude extension to higher dimensionality (Peuquet, 1984). A notable exception is the concept of time in narratives that may run contrary to the single uniform path (Caquard & Fiset, 2014). The third category of the data cube captures observable attributes of objects or events that do not belong to either of the other two categories. Attribute dimensions fall under labels such as "descriptor data" (Peuquet, 1984) and "multivariate" (D. Guo et al., 2006). The number and qualities of dimensions in the attribute category are flexible. Breaking down the three categories of geospatial data into component dimensions come into play when discussing representations and the alignment of data dimensions with the display dimensions. Arranging data values according to spatial dimensions gives rise to the concept of spatial patterns.

From a geography perspective, interest lies not in individual data points, but in relationships between those points and variability among their respective values over geographic space. Robinson et al. (1995) point to a shift during the Enlightenment from depicting the locations of entities on reference maps to a "holistic concern with the spatial extent and variation of features" (p. 27). This holistic concern describes *spatial patterns* that emerge from the data when taken at an appropriate scale and are fundamental to spatial knowledge (Golledge, Marsh, & Battersby, 2008; Nystuen, 1968). Understanding the occurrence of phenomena across space is fundamental[1] to geographic inquiry (Robinson & Bryson, 1957), and researchers describe *spatial patterns* or *distributions* by the existence (or lack) of structure (Robinson, 1975), areal pattern (Dent, 1999), and change over time (P. J. Gersmehl & C. A. Gersmehl, 2007). These descriptors do not quantify spatial patterns, but describe characteristics that are of interest when considering spatial patterns. Further, beyond a single instance of spatial pattern, comparison of distributions is a common task within geography (Robinson & Bryson, 1957) and

[1] While fundamental, the concept of distribution spans levels of complexity. For example, Golledge et al. (2008) describes distribution itself as "simple", while *identifying arrangement of a distribution* and *measuring similarity of distributions* are classified as difficult and complex, respectively. This distinction between recognizing a phenomenon and reasoning about it is consistent with theories of development, in which the ability to perceive phenomena precedes the ability to create an external representation of them (e.g., Piaget & Inhelder, 1948/1956).

a subject of importance across many disciplines (e.g., Boyle et al., 2016; Mowrer, 1938; Nettl, 1960; Saltré, Duputié, Gaucherel, & Chuine, 2015). The research presented in this manuscript does not seek a measure of relative similarity between spatial patterns, but offers comparison as an example task to which geospatial data could serve as input. Moving forward, this chapter posits that spatial pattern is a concept and subject of sufficient consequence to warrant investigation of its representation. Spatial pattern defines the the scope of the discussion of representation and communication of geospatial data.

Two dimensions of spatial data and one dimension of attribute data are both necessary and sufficient to manifest a spatial pattern. While spatial patterns in real world data are time dependent, a snapshot of geospatial data can be taken at a single point in time. Data from a single point in time define a spatial pattern, and the specific time at which a snapshot records data provides context to help interpret or make meaning of such a pattern. In this research, three geospatial data dimensions (two spatial and one attribute) are the critical contributors to spatial pattern and multiple forms can express these thematic patterns.

Representations

Geospatial data take many forms, and these representations are intertwined with the meaning that can be made from the data and the uses to which they can efficiently be put. Importantly, any representation of geospatial data that is not the earth itself is a model, which is necessarily incomplete and inaccurate. A representation influences how we think about the data (see, e.g., Couclelis, 1992; Kitchin & Dodge, 2007). The conceptual framework organizes geospatial data dimensions across three categories (space, time, and attribute) and provides a useful reference frame to also describe representations of geospatial data in various forms. Common forms of geospatial data include machine readable digital data, mental representations, and geographic map artifacts. Each of these forms is distinguishable from the others by the medium in which it manifests. After expanding on the idea of alignment between the dimensions of geospatial data and those of its representation, this section addresses each of these forms in turn, as well as how they relate to one another. The discussion outlines how each representation relates to the dimensions of the data.

Effective alignment between the dimensions of geospatial data (as they exist in reality) and those of a representation influences the effectiveness of a representation. Such alignment considers both the number and quality of the dimensions, and assignments that fall within conceptual categories (Dubus & Bresin, 2013, e.g., mapping temporal data to the time dimension of a display) may provide advantages over those that cross conceptual boundaries.[2] The display must have a sufficient number of dimensions to represent all pertinent dimensions of the knowledge (Bertin, 1967/2011), possibly with redundant encoding. The identification of pertinent dimensions is tied to to the intended purpose or use of the representation (e.g., "functional context," Robinson, 1975). With an interest in spatial pattern, two spatial dimensions and one attribute dimension constitute pertinent dimensions of the geospatial data within the context of this research. Alignment between the qualities of the dimensions considers level of measurement and associated perceptual qualities. Qualities of the dimensions must be compatible, but are not necessarily equal; crafting representations that create

[2] The potential for within category symbolization to provide advantages does not deny potential advantages of crossing category boundaries. For example, visual displays arrange attribute data on spatial dimensions of the display to help identify multidimensional autocorrelation (Skupin & Fabrikant, 2003).

sufficient alignment between geospatial data and its representation is both an art and a science.

Machine Readable Digital Data

Modern systems that store, process, and archive data on computers use digital data formats.[3] Geospatial data is no exception, with active areas of research and development in spatial data structures, spatial query algorithms, and metadata standards to accompany spatial data. Within a digital data representation, a process of abstraction converts geospatial data to a series of ones and zeros (binary digits, or "bits"). Sources that create a digital representation and applications that interpret it share an understanding of the abstraction process (e.g., marshalling data) and such an understanding embeds meaning behind particular combinations of bits. While the spatial and temporal dimensions of the data, for example, do not necessarily reside at corresponding spatial locations on a computer disk and exist continuously and concurrently regardless of the time they represent, a known abstraction technique allows users of the data to infer meaning. As an abstraction, however, digital representations of geospatial data are incomplete. Theoretical computer science provides techniques for determining the minimum number of bits or information required to represent critical details and optimization may omit or simplify details that are not deemed to be critical. Despite the apparent objective nature of digital data, the choice of variables to record, the precision with which measurements are taken, researcher bias or expectations of the outcome all play a role in the completeness of a data set. For example, field recordings

[3] Digital data may be a record of primary source geospatial data, or an archival record of output from map design or data analysis (Donkin, 1970; Tobler, 1959). While recognizing the importance of other non-digital data records, such as handwritten field notes, this discussion restricts the scope of data to computer-based digital data.

are limited snapshots of a phenomenon or an event that can be interpolated with varying degrees of feasibility and accuracy (e.g., a discrete number of sensors at weather stations report temperature values), and software libraries optimize storage and processing algorithms to suit intended use cases (e.g., rendering visual display: digital objects for the management of spatial data within the GeoTools[4] are integrated with the Java Swing[5] graphics library). Even given the limitations associated with incomplete and subjective data, digital representations of geospatial data have great advantages. Digital data contribute to widespread availability of digital datasets, enhance feasibility of simulation studies and complex statistical analysis, and facilitate automation of both data analysis and display in geographic map artifacts.

The implementation of digital representations combine both artistic and scientific techniques. Design of digital representations benefits from a creative approach. With a plethora of choices in design, creative use of metaphor (e.g., philosophy of data structures informed by theories of spatial cognition, Couclelis, 1992) strengthens the alignment between dimensions of geospatial data and abstract dimensions of digital data, and thus increase fidelity of the representation to the underlying data. And scientific methods measure and assess the efficiency of resulting implementations. Multiple approaches that provide equivalent function can be compared to guide choices between implementations, selecting those that appropriately balance storage space and computational power as required by the application, which can be unique to geospatial data. Implicit authority of digital data to serve as evidence of “truth” means that implications of the choices in

[4] Open Geospatial Consortium, GeoTools, http://www.geotools.org

[5] Oracle, Java Swing.

implementing a digital representation of geospatial data are broad, and underscore the importance of provenance and reproducibility. Notably, however, an example from anthropologists and sound artists illustrates tension between varying levels of regard for digital data products based on their role as archival records in contrast to artifacts that were produced through conscious planning and design (Samuels, Meintjes, Ochoa, & Porcello, 2010). The value of digital data is not simply in their existence, but in the extent to which they support humans in making meaning or enhancing knowledge of the real world phenomena or events that they represent.

Mental Representations

Receiving input from first hand observations of the world or from digital data, the human brain (and those of other species) have an ability to encode, reason about, and remember geospatial data. The mechanisms that support mental representations are well understood for some aspects of perception and cognition, but are still largely theoretical for others. For example, the spatial locations of stimuli seen within the environment are represented by neurons in primary sensory cortices that have a similar spatial arrangement[6] (e.g., retinotopic organization of neurons, Constantinidis & Wang, 2004) and the location of neural activation of grid cells in the hippocampus corresponds with location in the real world (Hartley, Lever, Burgess, & O'Keefe, 2013). However, ongoing research is investigating the

[6] The observation of similarity between the spatial arrangement of environmental features and that of neurons which selectively respond to those features does not imply semantic importance. For example, a linear array of lights may be encoded as an aggregate attribute (e.g., angle of the line) rather than a persistent explicit representation of the location of each light as it ascends the visual processing pathways of the brain. Connections between lower and higher levels of processing are an an active area of research (e.g., Silson, Chan, Reynolds, Kravitz, & Baker, 2015), and while lower level perception in the primary sensory cortices is necessary to influence higher level cognition (MacEachren, 1995), it is not sufficient to reliably induce or predict a specific higher level response.

role of such organization at higher levels of cognition, such as the precise nature of mental representations of space (e.g., Chrastil & Warren, 2014; Montello, 1997) and the brain structures that support them (e.g., Boccia, Nemmi, & Guariglia, 2014). And alignment of the time and attribute dimensions of geospatial data with respective dimensions of a mental representation are inconsistent or unknown; for example, a gradient of frequency selective neurons (an attribute of sound) are arranged spatially in the cortex. Along with wide acceptance that the brain stores a representation of geospatial data (at whatever level of alignment between the data dimensions and the dimensions of the mental representation) researchers accept that mental representations are incomplete subsets of reality (Wood, 1972). Processes that support perception encode details to address a given task, while unnecessary detail is ignored. And the brain fills in details that are needed to reason through a logical task but may be missing. Not only are mental representations incomplete versions of geospatial data (Montello, 2002; Tversky, 1993), but their formation is highly variable. For example, empirical studies reveal individual differences in map reading (e.g., Hegarty, Montello, Richardson, Ishikawa, & Lovelace, 2006; Thorndyke & Stasz, 1980). Previous experience influences mental representations (Montello, 2002; Muehrcke, 1973) and uptake of information from a geographic map (Schito & Fabrikant, 2018). Further, encoding strategy is flexible (Kulhavy & Stock, 1996; Nees & Walker, 2008, 2011) and attention modulates the selection of details to encode. While incomplete mental representations lead to errors in spatial judgements (Golledge, 2002), they are sufficient to support many purposes, even being noted as an "accomplished system for the representation, explanation, and prediction of geographic phenomena" (Couclelis, 1992).

Mental representations, like digital representations, are understood through both creative and scientific approaches. Until relatively recently inspection of the brain in living humans was largely limited to observable behaviors in response to the presentation of controlled stimuli. Despite and perhaps because of those limits, psychologists develop elegant experimental designs to isolate phenomena of interest and explain observable behaviors. And pioneering work in spatial cognition advances knowledge of how the brain acquires and represents information (e.g., O'Keefe, 2014; O'Keefe & Dostrovsky, 1971). Theories of spatial knowledge and mental representations of space (e.g., anchor point theory, Couclelis, Golledge, Gale, and Tobler, 1987; cognitive collages, Tversky, 1993) emerge and evolve as supporting or contradictory evidence emerges. As technologies progress, additional tools for the scientific investigation of mental representations provide insight into both the influence of external stimuli and the manifestation of mental representations in expressed behaviors. A mixture of creative and systematic approaches to understanding mental representations of geospatial data give the power to capture nuances of the human information processing (in contrast to the deterministic information processing on computers). And understanding mental representations allows cartographers to apply theories of spatial cognition to map design (e.g., "cognitive cartography", Montello, 2002), educators to design effective tools to teach spatial thinking (P. J. Gersmehl, 2011; Golledge et al., 2008), and researchers to interpret geographic map artifacts as reflections of the people and situations that produced them (Kitchin & Dodge, 2007). To both create and externalize (e.g., sketch maps Newcombe, 1985) mental representations, people employ geographic map artifacts.

Geographic Map Artifacts

Geographic map artifacts – and thematic maps in particular (Figure 1), which support investigation of relationships between data at different locations (Montello, 2002; Robinson et al., 1995) – serve as communication channels (Cobb, 1977; Koláčný, 1969; Morrison, 1974), data storage volumes (Cobb, 1977; Donkin, 1970; Morrison, 1974), and persuasive instruments (Kitchin & Dodge, 2007). In the context of this research, communication of spatial patterns from digital data representations in thematic geographic maps is the primary objective. Like machine readable data, geographic map artifacts have a physical form, but, instead of targeting use on and by machines, geographic maps are designed to be perceived by humans, and symbolization influences the way a map observer perceives and reasons about the geospatial information contained in a map (Golledge, 2002; Ogao & Kraak, 2001). The geographic map artifacts themselves have dimensionality that is inextricably tied to the modality of the display. Symbology assigns dimensions of the data to dimensions of the display, and the modality of the display determines the available display dimensions. Through alignment of data dimensions with display dimensions, geographic map artifacts depict generalized and simplified geospatial data, and decisions that lead to the abstract representation are influenced by the intended use (for a review, see Medyckyj-Scott and Blades, 1991; Robinson et al., 1995). The use of metaphor (Fabrikant, Montello, Ruocco, & Middleton, 2004; MacEachren et al., 1999) and structure (Gattis, 2001) facilitate communication of a spatial pattern despite incomplete data, but a map documents only the map maker's interpretation of their physical, social, or cultural environment. As a design process, one set of geospatial data gives

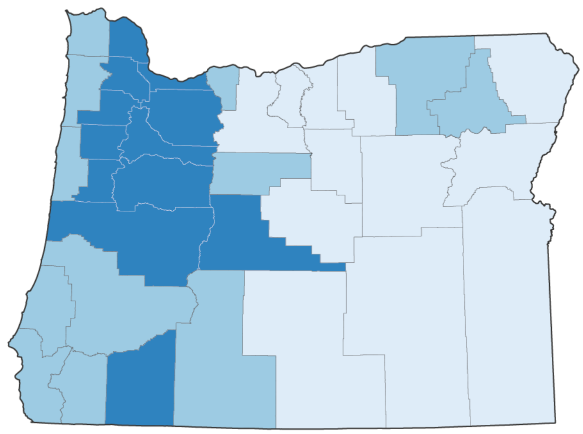

FIGURE 1. A choropleth map, in which color represents population density, is one example of a thematic map that uses a visual cartographic technique to illustrate spatial pattern. *Data Source: U.S. Census Bureau*

rise to many possible geographic map artifacts (Robinson, 1975), and the result reflects the cartographers subjectivity (Montello, 2002).

While empirical studies provide some guidelines, many of the design choices are the result of the cartographers intuition or preference (Montello, 2002; Muehrcke, 1973). In the case of geographic map artifacts that serve as a communication medium or evoke and support thinking about spatial patterns, these imperfections and subjectivities mirror those of the target mental representations. A common understanding of the symbolic representation, which the map maker and the map reader share, embeds meaning in the features of a geographic map artifact without requiring absolute accuracy. And incomplete data are not necessarily problematic; geographic map artifacts target the human perceptual systems, which gather only a partial view of the world. Instead, the symbology is an abstract representation that must maintain high fidelity to the essence of the data (P. J. Gersmehl, 1985). And, in the case of thematic maps, conveying the gist of the data and spatial patterns of relationships between data at different locations has higher precedence (Robinson, 1986) than query for individual data values.

The design of symbology demonstrates the art and science of geographic map artifacts, which require a systematic relationship between the data and the display. There is no single optimal design (Kitchin & Dodge, 2007) and existing guidelines are a mixture of anecdotal experience (Couclelis, 1992; Muehrcke, 1973) and empirical evidence from psychophysics (Montello, 2002). While geographic maps embrace social and artistic expression to varying degrees (e.g., see discussion in MacEachren, 1995), all design processes involve decisions that are not fully specified by the requirements. With an exponential number of possible pairings, aligning data variables with display dimensions is an artistic part of design (noted in the context of sonification design Asquith, 2013), but computers can alleviate the time cost of exploring multiple different arrangements (Essinger, 1986). Since the late 1950s (Montello, 2002) and gaining momentum in the 1970s (e.g., Morrison, 1974; Muehrcke, 1973) cartographers' interest in automation lead to work formalizing design knowledge and establishing deterministic guidelines. The interest in automation offers one explanation for subsequent resurgent interest studies of the psychophysical properties of map symbols (Montello, 2002). Computers support map design by providing functionality to both apply rules that systematically translate data into display dimensions and subjectively implement exceptions to those rules. For example, workflows for visual geographic maps apply both systematic assignment of data values to features in the map display (e.g., automatically mapping ratio values to graduated symbol sizes in ArcMap) and flexibility to adjust results of the automated process (e.g., manually[7] increasing the distance between symbols to avoid occlusion in Adobe Illustrator).

Relationships Between Representations

Digital, mental, and geographic map artifact representations do not exist in isolation; one representation may serve as input to or output from another. The choice of a representation influences efficiency of later use and translation to an alternate form. And the influence of a data model propagates downstream through the communication channel, such as a geographic map artifact, to interact with efficiencies in use of the data by map readers. For example, theories of spatial cognition influence digital representations (e.g., raster and vector formats in GIS, see Couclelis, 1992). And digital representations of spatial data that optimize data structures for graphic rendering discard data that may be critical to alternative display modalities (e.g., screen readers Mynatt & Edwards, 1992).

Each translation to, from, or between representations loses data. For example, the contents of a digital data set filter the physical world into discrete observations. A geographic map artifact that displays such data sets cannot depict values that were never recorded (but, notably, interpolation techniques mitigate the impact of sparse data sets with varying degrees of accuracy). Further, the map reader cannot extract meaning from the map that is depicted in a way that is not both perceptible and meaningful. The cartographer's experience, positionality, and a priori expectations influence selection and symbolization of data. For example, in this research, pilot study data reveal the use of learned strategies for exploration

[7] While some adjustments can be automated, e.g., dynamic label placement, manual efforts to polish map layout provide an added value in map design (Brewer, 2016).

of two dimensional data space, “transects”, and use of contour lines to draw the surface that they hear.

Similarly, translation of one geographic map artifact into another often loses data. Translation of a graphical representation to other display modalities (an “adaptation strategy”, Savidis and Stephanidis, 1995, widespread in applications that address accessibility post hoc) may produce suboptimal results (Frascara & Takach, 1993; Frauenberger et al., 2005; Ojala, Lahtinen, & Hirn, 2016; Rice, Jacobson, Golledge, & Jones, 2005; Savidis & Stephanidis, 1995). The specific data that is lost depends on the form of the representation and the processes that produced them.

Just as translations between representations may lose data, eliminating unnecessary translations between representations preserves data. Ideally this would mean a mechanism to insert a digital data representation directly into a mental representation (“open brain, insert map here”). But short of direction injection, a perceptible geographic map artifact is necessary. Design of auditory geographic maps in this study originates with digital data and explicitly avoids an intermediate visual geographic map artifacts. As a display medium which has a limited presence in cartography, this means first investigating how geospatial data can be represented in sound. The next section explores sound as a display modality, both in general and as a medium for the display of geographic map artifacts.

Sonification and Audio Cartography

Sonification is the representation of data in non-speech audio (although the definition has a history of varying details and nuances, de Campo, 2007;

Dubus and Bresin, 2013; Supper, 2015). To represent data,[8] sonification designers create a relationship between data values and dimensions of non-speech audio.[9] Parameters of sound synthesis modulate the physical properties of sound waves and provide some control over perceptual characteristics of the result, although still not completely predictable. As an introduction, this section provides an overview of several aspects of sonification that inform the design of auditory geographic map artifacts (see additional review and history of sonification in Hermann et al., 2011; Kramer et al., 1997) and reviews previous work on auditory geographic map design and evaluation.

Dimensions of Audio

Audio is characterized by both properties of a physical sound wave and by the qualities perceived through the human sense of hearing. Sound waves are longitudinal pressure waves whose physical properties include frequency, or the number of oscillations per second (hertz, Hz), and amplitude, or the magnitude of displacement from a reference level (decibel, dB). When sensed by the ear, sound waves produce perceptual sensations that are associated with, e.g., pitch and loudness. Relationships between physical properties of a sound wave and the perceptual properties that they evoke (see also reviews of psychoacoustics,

[8] Sound can also evoke emotion, which has been explored in cartography (e.g., Edsall, 2011), but such applications are distinct from the display of spatial patterns in geospatial data and beyond the scope of this manuscript.

[9] Speech-based audio plays an important role in in rich oral traditions and in some computer-based applications, but is beyond the scope of this work. For example, Aboriginal songlines relate spatial information. And turn-by-turn directions delivered through synthesized speech are a widely used feature in mobile device applications (e.g., talking GPS, see review in Currier, 2011) While a turn-by-turn sequence can support navigation along a linear path, such an approach does not directly generalize to thematic data. Verbal descriptions of thematic data are under investigation (Frye, 2015, e.g., GeoDescriber). Natural language to describe spatial patterns, however, is often inefficient (Rinck, 2005) and relies heavily on interpretation by the describer.

e.g., Deatherage, 1972) mean that sound can serve as a channel to communicate data, including multi-dimensional data. However, these relationships are neither deterministic nor independent (Brewster, 2003). For example, frequency of a sound wave is often perceived as pitch, but may be influenced by the presence of harmonics. And amplitude loosely translates to loudness, but the perceived loudness is affected by the frequency of the sound. Recognizing that the plethora of perceptual dimensions of sound are more numerous than can be exhaustively covered here (Levitin, 2002, identifies six to eight auditory dimensions that can be perceived and remembered; Krygier, 1994, describes nine "abstract sound variables"; Dubus and Bresin, 2013, list thirty auditory dimensions across five categories for use in parameter mapping sonification), this section presents selected dimensions of sound and perceptual qualities, with which the physical properties of the sound wave are strongly associated (i.e., physical properties of the sound wave can be modulated to encode and communicate data). The selected dimensions of sound are commonly used in sonification and are detectable by untrained listeners.[10]

[10] In this usage, "untrained" means that detection of the audio characteristic does not require specialized formal training (e.g., ear training in music; this definition does not preclude a need for training to associate a specific audio characteristic with its semantic meaning in the map). Further, musical training does not necessarily advantage listeners in the identification of sound events (Alexander, O'Modhrain, Gilbert, Zurbuchen, & Simoni, 2012), although training may provide vocabulary to describe what they hear. An analogous example in visualization is the detectability of different hues; a sighted person can determine that a blue and a red map symbol have different hue without special training to view graphics. While Schito and Fabrikant (2018) report statistically significant correlations between previous experience and map reading performance, the influence of music experience does not stand out from that of domain expertise ("terrain interpretation") or familiarity with computers ("technical ability").

Frequency

Mapping data values to frequency is common in sonification and software functionality to implement this mapping is readily available. Pitch is the perceptual quality strongly associated with frequency (although there is an interaction between frequency and perceived loudness). Frequency readily communicates ordinal values (Krygier, 1994), supporting the detection of relative differences. Collections of frequencies create patterns in two ways. The spectral shape of harmonics, or frequencies that are integer multiples of a common lowest frequency, contribute to the timbre of a sound. Groups of notes that have varying fundamental frequencies create musical intervals and chords, whose detection and interpretation is culturally informed (informal training) and identification and naming can be learned (music training).

Amplitude

Software functionality to control the amplitude of a sound wave is widely available both during initial waveform rendering and in post-processing. Amplitude relates to perceived loudness (Deatherage, 1972), and depends on the frequency of the sound. Published tuning curves, e.g., provide quantitative scaling factors to achieve "constant loudness" by attenuating amplitudes based on frequency. Amplitude values can convey ordinal data (Krygier, 1994), but are limited to relative comparisons. Amplitude-based patterns may occur over time, but are less apparent when sounds occur simultaneously. Sound waves with low amplitudes may contribute to an overall quality of the sound (timbre), but be masked by stronger sounds.

Spatial Location

The human auditory system localizes sound sources using multiple monaural (e.g., amplitude) and binaural cues (e.g., interaural time difference). Perceived azimuth (horizontal, or direction in the axial plane) has relatively high resolution, while distance and elevation (vertical, or direction in the sagittal plane) have lower resolution. Notably, localization of sounds is necessarily egocentric, or relative to the listener; conveying allocentric relationships between data requires a perspective transformation by the listener. Despite the apparent simplicity of using spatial dimensions of the display to represent spatial dimensions of the data, effective use of the spatial display dimensions is unresolved (Nees & Walker, 2009). Guidelines do recommend the use of spatial sound to draw attention to a single location or variable within a display (L. Brown et al., 2003; Krygier, 1994) or encourage an immersive experience (Schito & Fabrikant, 2018).

Note Rate

Many physical properties of a sound wave can contribute to perceptual grouping of sound events that occur over time. In this discussion, the note rate describes both the duration of individual sound events and the inter-onset interval that separate sound events in time (two physical characteristics of sound that may be considered independently. Co-varying duration and onset of sound events creates an audio symbol that unfolds over time, but has a persistent perceived quality of fast or slow.[11] Relative differences in the note rate (e.g., “faster”) convey ordinal relationships in the underlying data. The note rate is a periodic pattern of sound

[11] The perception of note rate requires sufficient time to elapse in order to detect rate. What constitutes sufficient time depends on the duration of sound events. A similar property is also true of frequency. If a sound event is too short, human hearing detects a click devoid of pitch.

events. In this use of periodic note rate, the temporal aspect of the display is subservient to the persistent quality of fast or slow.[12]

<u>Pattern</u>

> *"But hearing data, it turns out, also can open new scientific frontiers. That's thanks to the remarkable human ability to parse sound for patterns and meaning. 'The auditory system is the best pattern-recognition device that we know of,' says Bruce Walker, a professor of psychology and director of the Georgia Institute of Technologys Sonification Lab."* (Hadhazy, 2014)

Multiple characteristics of sound within and across dimensions of audio combine to create patterns. Patterns emerge not from individual entities, but from relationships between multiple characteristics or instances of sound events when taken at an appropriate scale. For example, multiple frequencies that occur simultaneously produce chords (intervals between frequencies conform to musical structure). And rhythmic combinations of notes can produce emergent patterns (e.g., streaming in African xylophone music Bregman, 1990). Or combinations of audio characteristics create sounds that have unique or identifying qualities (e.g., timbre of sounds associated with the physical materials that produce them Paté, Boschi, Carrou, & Holtzman, 2016). For example, assignment of multidimensional data sets to audio dimensions produces "hiss" or "tuning-fork-like" of qualities when sonifying nine data dimensions (Yeung, 1980). Perceived as qualities of the sound (e.g., hiss or fast note rate), pattern aligns with the attribute category of

[12] Interactive interfaces, such as those that represent data values according to a cursor location, movement speed confounds the perception of spatial patterns and temporal patterns; variability in the speed as a cursor moves across regularly spaced data features can produce an irregular temporal pattern that may be misinterpreted as spatial variability (Brittell, Young, & Lobben, 2013). In the use of note rate, the timeline of the sonification is controlled (i.e., not interactive) to avoid unintended emergent temporal patterns.

the geospatial data cube. Design that incorporates some common patterns, such as musical chords, are straight forward to implement. And exploratory sonification leads to emergent patterns (de Campo, 2007). However, complex patterns, such as those in African xylophone music, require a high degree of skill in composition and are less amenable to systematic production from diverse data sets. The ability to communicate of pattern motivates exploration of sonification to represent complex data sets (Alexander, Zurbuchen, Gilbert, Lepri, & Raines, 2010; Supper, 2014; Walker & Mauney, 2010).

The dimensions described above by no means constitute an exhaustive list. Nor do they represent a minimum list. For example, any perceived sound wave has a spectral envelope that is often associated with timbre, or the quality of a sound that lets the listener distinguish between human voices or musical instruments. These dimensions do, however, provide a sufficient base in which to ground discussion of the design of auditory geographic maps.

Parameter Mapping Sonification

Having identified dimensions of data and dimensions of auditory display, it remains to explore mappings between data and display. Parameter mapping sonification (akin to the term "map symbology", from cartography) creates a systematic relationship between data values and display dimensions, and is a popular sonification technique (Dubus & Bresin, 2013). But no prescriptive guidelines apply to all design situations (Nees & Walker, 2009; Robare & Forlizzi, 2009). In the absence of prescriptive design guidelines, or in the case of audio cartography a general lack of design guidelines, similarity to sounds experienced in the real world and guidelines from the design of auditory computer interfaces help guide choices in parameter mapping.

Parameter mappings that are "intuitive" or "natural" reflect sounds with which we have experience in the real world (Dubus & Bresin, 2013; Gaver, 1989). Familiarity reduces the time required to learn a parameter mapping (Dingler, Lindsay, & Walker, 2008) and decreases susceptibility to ambiguity (Dubus & Bresin, 2013). For example, sonification of population data might start with an audio clip of the sound of voices in a crowd and modulate frequency or amplitude to encode data values. However, as an abstract representation, sonification in the applied domain of geographic maps may represent data for phenomena that do not emit sound (Montello, 2002), or exceed what could be experienced first hand.

The spatial extent of map contents, or the geographic space represented in a map, may far exceed the the bounds of our perceptual systems. We can explore a map of the earth, which stretches 24,901 miles along the equator (Evers, J. (Ed.), 2011), even though we can only feel items within arms reach (about three feet), see the horizon at a distance of just over three miles (French, 1982), or hear a low frequency foghorn, which was designed for long distance perceptibility, up to eight nautical miles (9.2 miles) away (Clingan, 2017). A desire to represent in maps that which we understand to exist, but is beyond what we can directly perceive leads to abstract symbolic representations in geographic maps. Further, a distinction between what we can perceive first hand from our environment and what a map can represent reinforces the geographic map as a tool for abstract representation. A strict alignment between categories or dimensions of geospatial data and a geographic map is not necessary: the spatial dimension of the data need not align with the spatial dimension of the display.

Where mapping to natural sounds fall short, a growing body of literature provides guidelines for the use of audio in computer interfaces (e.g., Brewster et al., 1995; Frauenberger et al., 2005; Nees & Walker, 2009; Sumikawa, 1985) and data displays (e.g., L. Brown et al., 2003; Flowers, 2005; Walker & Kramer, 2006). General guidelines recommend the use of metaphor. The "distance-similarity" metaphor (Fabrikant et al., 2004; MacEachren et al., 1999), for example, may partially explain the effectiveness of spatialization, or presentation of non-spatial data on the spatial dimensions of a display. And explicit structure in the display may reflect inherent structure in the data to increase perceived similarity (Gattis, 2001). For example, shared linear structure makes natural language a good display option for route descriptions (Richter & Winter, 2014).

Auditory Display Applications

Auditory displays have found use in data exploration by research scientists[13] (e.g., Alexander et al., 2010; Cherston, Hill, Goldfarb, & Paradiso, 2016; Diaz-Merced et al., 2011; Hegg, Middleton, Robertson, & Kennedy, 2018; St. George, Crawford, Reubold, & Giorgi, 2017) as part of research communication with the public (e.g., Alexander et al., 2010; Asquith, 2013; Supper, 2014), and in applications that improve access to data for people who are blind (e.g,. Zhao et al., 2008). Two things that these diverse applications have in common are the qualities of the underlying data and presentation in an auditory display. These sonification projects all create representations of large, multidimensional data sets in audio.

[13] Amid increasing appreciation of the power and possibility of auditory display across multiple disciplines and settings (Alexander et al., 2010; Kramer et al., 1997), however sonification continues to contend for legitimacy in scientific circles (Asquith, 2013; Samuels et al., 2010; Supper, 2015).

Existing cartographic designs occasionally include an auditory display, yet tend to be multimodal displays that rely on a visual (see review in Brauen, 2014), proprioceptive (e.g., Brittell et al., 2013), or haptic (e.g., De Felice, Renna, Attolico, & Distante, 2007) components to communicate location. Early sound synthesis tools offered a promising new technique for cartographic design (e.g., Fisher, 1994; Krygier, 1994 augment visual geographic maps with audio, and Cassettari and Parsons, 1993 consider audio as an emerging new data type). After an intervening decline in interest, the introduction of browser-based audio synthesis capabilities renews interest in audio to augment interactive web maps (see Brauen, 2014, for review). However, these displays typically rely on visual information to convey spatial location. Accessibility applications also adopt auditory displays (both speech and non-speech), which commonly complement another modality (see also Currier, 2011). However, the interface through which map readers query the display is serial; map readers move a stylus around a tablet device, soliciting sonification of the single data point that corresponds to the location of the cursor. And listeners struggle to reconstruct general spatial information from the sequential display (Alty & Rigas, 2005; Brittell et al., 2013; Delogu et al., 2010). These displays cast audio to a supporting role augmenting or providing redundant encoding of a subset of the information already available in a visual display. While audio serves an important role in these multimodal displays, the ability of audio to convey pattern presents a relatively unexplored frontier in cartographic representations of the spatial dimensions of data.

Stand-alone auditory displays, which encode spatial location in audio, are more rare (see also Brittell, 2018). Several examples come from applications that display spatial data, but not necessarily *geo*-spatial data. Flowers, Buhman, and

Turnage (1997) uses frequency and time as two axes for an auditory display of scatterplots, depictions of spatial data that is not specifically geographic. And L. Brown et al. (2003) published guidelines for the design of auditory mathematical graphs that continue to rely on time as the x-axis. Among the limited number of auditory geographic maps that convey geographic location through audio, converting two dimensional spatial data into a linear sequence is the most common design strategy. A virtual cursor imitates the movement of a pointing device (but without the proprioceptive feedback). Examples include iSonic ("gist" action, Zhao et al., 2008), in which the virtual cursor follows a predetermined scan path, and AudioGraph (Rigas & Alty, 2005), in which paired auditory notes encode coordinate locations and a virtual cursor traces shape outlines.

Implementation of a sequential display of data values is straight forward, but there are usability issues in the resulting displays.

Toward Audio Cartography

Given this background, the objective of the research presented in this manuscript is to better understand how the temporal order of an auditory display influences map listening and learning. Guidelines from computer-based user interfaces and sonification in other application domains provide a starting point. Where guidelines fall short, there is a dearth of established convention to fall back on. Or, rather, entrenched cartographic convention is based on visual displays, which stand to bias or hinder design decisions (Frauenberger et al., 2005; Nees & Walker, 2009). Instead of falling back on visual traditions, there is a need to step outside our comfort zone and better understand how to represent geospatial data in audio, including the implications of transforming spatial data dimensions into linear sequences.

Sonification of geospatial data offers benefits in data exploration and accessibility. Novel approaches to data display (alone or in multimodal applications, Cartwright, 1996) have the potential to expose patterns in the data that are not apparent in alternate displays (Diaz-Merced et al., 2011). And a better understanding of how to communicate geospatial data through audio can lead to alternative display tools, reducing the need to seek exceptions to social and legal accessibility requirements citing, e.g., the "visual nature" of GIS tools (C. Brown & McKinney, 2015).

Audio synthesis technologies are widely available and auditory displays have no inherent deficiency that would preclude use in geographic map design; the limitation is a lack of prior knowledge about how to wield the tools of auditory display to sonify geospatial data. Even more generally, sonification is a relatively new technique. Advancing from a debut of computer music in 1951 to widely available sound synthesis in the 1980s (Robare & Forlizzi, 2009), technology to support auditory display is now widely available. But guidelines to effectively use sonification are not yet mature (Nees & Walker, 2009). And, in any application domain, the adoption of new tools disrupts the status quo and pushes designers outside of their comfort zone. For example, Sigsworth (2018) notes that with the the arrival of electronic drum machines imagination poses the main limitation; after spending years honing their skill at controlling physical drumsticks he observes how it was "their minds and not their limbs that were limiting them." Although, experience with cartographic techniques (or with drumsticks) provides invaluable intuition to overcome gaps in formal design guidelines (Montello, 2002; Muehrcke, 1973), a lack of guidelines for auditory-geographic-map design partly because auditory-geographic-map design hasn't widely been done before.

CHAPTER III

METHODOLOGY

An empirical study explores the influence of temporal arrangement of data within an auditory display and provides evidence to address the two research questions. Three auditory map types demonstrate varied temporal arrangement of data within the audio stream. The three auditory map types encode location in time, in acoustic properties of sound, or in both time and acoustic properties. The encoding of location in time mimics a previously established approach to map sonification, while the encoding of location in two acoustic properties of sound (frequency and note rate) is a novel approach. Evaluation follows a two-pronged approach to understand the contribution that the temporal arrangement of an auditory display makes to the communication of spatial information through an auditory map. Borrowing a model from psychology, behavioral and functional magnetic resonance imaging (fMRI) methods offer complementary perspectives on the relationship between stimuli and perception[1] (Goldstein, 2014). The behavioral data inform the investigation of a stimulus-perception relationship, addressing the first research question:

RQ1: How does the temporal aspect of auditory map symbolization influence effectiveness in communicating general spatial patterns in the data?

[1] In this use, "perception" encompasses detecting, encoding, and recognizing information from the environment. This usage contrasts with discussion of perception in the cartography literature that often focuses on psychophysics and low-level stimulus processing (Montello, 2002).

while the neuroimaging data offer insight into the stimulus-physiology relationship, addressing the second research question:

RQ2: How does neural activation in response to serial audio symbolization of a geographic map contrast with that of simultaneous audio symbolization?

Rather than relying solely on theoretical and established (often ocular-centric) ideas about cartographic design, this research adopts a methodology that combines behavioral and neuroimaging metrics to investigate these research questions. The combination of behavioral and neuroimaging methods offers insight into the implications of map design beyond that available from either method alone. Illustrative instances of auditory symbology demonstrate feasibility of the implementation and enable assessment. Guidelines from research in human-computer interaction, theories of spatial information from geography and spatial cognition, and existing auditory maps inform the design. Yet a scarcity of established guidelines for usable audio-only map displays necessitates the introduction of novel designs. Behavioral data provide a measure of map effectiveness, and demonstrate the degree of correlation between systematic modulation of the stimuli and variations in performance. Behavioral geography has widely adopted behavioral methods (Montello, 2016), which have been long established in psychology. The neuroimaging data provide a measure of relative brain activation in the context of a map reading task. Following technological and methodological advances in cognitive neuroscience, geographers have relatively recently started to probe patterns of brain activation to explain observed differences in map effectiveness (P. J. Gersmehl & C. A. Gersmehl, 2006; Lobben et al., 2009). Augmenting behavioral data with evidence of neural responses to map stimuli moves beyond a question of whether or not a specific design "works" to investigate the mechanisms behind *how* or *why* aspects of a design work.

This chapter describes the design of three auditory geographic map types and the methodology employed to evaluate differences between them. First, the text introduces several factors that influence the design of auditory symbology and the resulting auditory geographic map artifacts. The thematic geographic maps used in this study are auditory representations of geospatial data, and the symbology strives to isolate the temporal component of the display. The second section describes the methodology and test instrument used to evaluate the auditory map types. Chapter V: Discussion describes the validity of this approach in greater detail.

Design of Auditory Geographic Maps

Geographic map designs realize a balance between precise display of the underlying geospatial data, emergent qualities of the rendered map artifact, and constraints that the research design imposes. The design of auditory geographic maps are no exception. This section describes the data dimensions selected for representation in the auditory display, two digital representations of those data dimensions, the parameter mappings selected to render three auditory geographic map types, and details of the experimental design with which the design choices were intertwined.

Geospatial Data Dimensions and Measurement Level

With the objective of representing and communicating spatial patterns in geospatial data, the auditory maps symbolize three dimensions of the underlying data. Two spatial dimensions describe location within a planar raster. And a single attribute dimension encodes (synthesized) population data values. The two spatial dimensions are orthogonal and establish a Cartesian plane.[2] Eight uniform discrete units subdivide each axis, creating an eight-by-eight raster grid.[3] Both the grid cells and the overall map extent are isotropic, which enforces an objective equality (but ignores possible subjective or perceptual differences, e.g., horizontal vs. vertical distance judgements in both auditory and visual perception). The attribute dimension consists of population data (see the description of the stimuli below). For simplicity, three ordinal levels aggregate numeric population values.[4] These data are recorded in digital files that are used as input to automated processing that renders auditory geographic maps.

[2] Geographic projections systematically flatten the earth's three dimensional surface accompany the extensive display of geographic maps on flat surfaces such as paper and computer screens. In auditory geographic maps, the need to project geospatial data into a flat plane is not obvious and even reinforces the primacy of a graphical representation (an inferior approach to auditory map design, as argued in Chapter II: Background). Increasing ability to simulate three-dimensional figures in computer graphics and adoption of auditory displays in cartography are changing the role of projections.

[3] In the design of visual geographic maps the cartographic process involves generalization and simplification of the geospatial data (Muehrcke, 1973; Robinson et al., 1995). Similarly, generalization and simplification processes are applied to cartographic design of non-visual maps, however, the degree of data reduction tends to be much greater in, e.g., tactile (Tatham, 1995) and auditory displays. The extreme simplicity of the data used in this study makes sequential sonification of each grid cell feasible based on both limits to working memory and reasonable duration of study sessions, but simplification to that degree is not believed to be strictly necessary for auditory display.

[4] The number of classes limits the resolution of the spatial pattern that a geographic map can convey, but has only a weak influence on the impression of spatial patterns (evaluated in the context of visual geographic maps, MacEachren, 1982a) Within reason, a larger number of classes could be encoded in a single display dimension; human perceptual abilities bound the extent to which a single auditory dimension can be subdivided and still effectively communicate data values (Pollack & Ficks, 1954).

Three Auditory Map Types

Two approaches to using time in auditory geographic map design explore the implications of temporal arrangement on efficacy of representing spatial location and conveying spatial pattern in audio. A *sequential* map type mimics existing auditory geographic map displays. Each cell of a raster data set is represented by a sound event, such as a musical note. A *concurrent* map type compresses the sound events in time, while avoiding auditory masking (Deatherage, 1972) from the simultaneous presentation of multiple sound events. Each cell of a raster data set is assigned a location with the sound design space; producing multiple sound events at the same time produces complex sounds. But substantial differences between these two map types confound attempts at direct comparison.

The sequential and concurrent map types differ in two important ways. First, the number of display dimensions used to encode two spatial dimensions of the data differ. The sequential map type reduces the two spatial dimensions to a single time dimension, and uses relative order of the sound events in time to encode location; a single sound event carries no location information. In contrast, the concurrent map type provides information about location within each group of sound events, which represent a single cell of the raster data. Notably, these two parameter mappings are not mutually exclusive. Second, the amount of information encoded in a single group of sound events differs. In the sequential symbology, each sound event explicitly encodes a single attribute value, devoid of location information. In contrast, a sound event in the concurrent symbology, encodes three data values in each sound event: two spatial dimensions and one audio dimension. Rather than a direct comparison of the sequential and concurrent map types, an additional hybrid map type creates a set of three map designs with incremental differences.

An *augmented-sequential* map type serves as a hybrid between the sequential and concurrent map types (Figure 2). Using both temporal order and frequency-note rate pairs, sound events in the augmented-sequential map type provide redundant location information. In contrast to the sequential map type, frequency and note rate augment each sound event with explicit location information. Compared with the concurrent map type, augmented-sequential map type uses the same set of possible (three parameter) sound events, but, the sound events occur sequentially in time. In this way, two paired comparisons tease apart the influence of explicitly encoding additional information in each sound event (sequential vs. augmented-sequential) from the effect of temporal order (augmented-sequential vs. concurrent).

These three map types represent only a small subset of the vast number of possible parameter mappings (see Nees & Walker, 2009, and discussion in Chapter II: Background). The chosen parameter mappings are neither optimal for use in more general audio cartography applications, nor are they the most aesthetically pleasing combinations. They do, however, satisfy the experimental constraints (see design constraints, described below), and support isolation of temporal arrangement for investigation. Beyond simply satisfying constraints, the selected parameter mappings are also amenable to automation. Automation encourages reproducibility and minimizes subjective decisions in the cartographic implementation of stimuli for experimental evaluation. The remainder of this section provides additional detail on the design and implementation of the three auditory geographic map types.

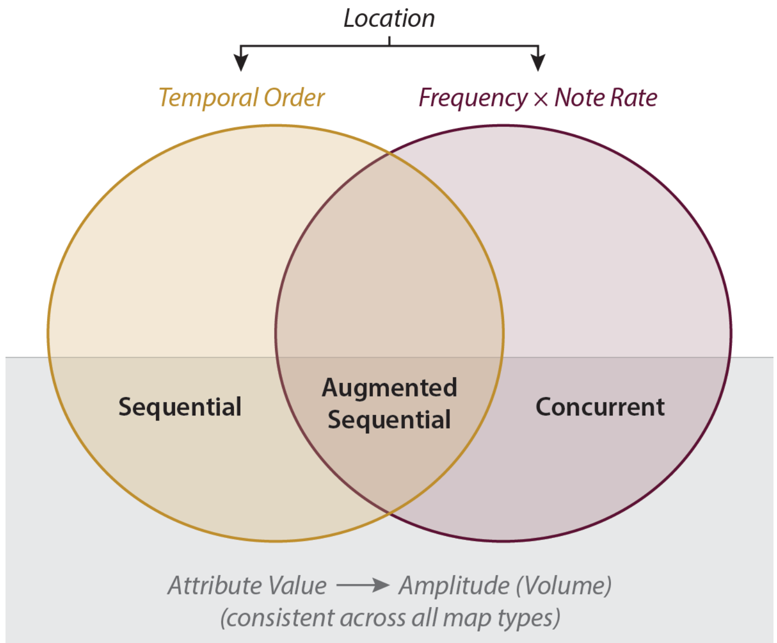

FIGURE 2. Three auditory geographic map types encode spatial location in audio. Two dimensional location is symbolized on a single time dimension of the display (sequential), two attribute dimensions of frequency and note duration (concurrent), or redundantly on both time and frequency-note duration pairs (augmented-sequential). Across all three auditory geographic map types, a single attribute dimension of the display (amplitude) encodes the attribute dimension of the geospatial data.

Digital Representations

The geospatial data and auditory geographic maps were stored in digital files with distinct formats. Two equivalent digital data representations support auditory map rendering under different uses of the time dimension of the display (Figure 3): the auditory geographic maps present spatial data either one data record at a time (sequential and augmented-sequential) or with multiple data values overlapping in time (concurrent). A matrix representation, which mirrors a common data format used to represent raster geospatial data, records a data value for each cell of the grid in a two dimensional array. A sequential audio rendering reads data values

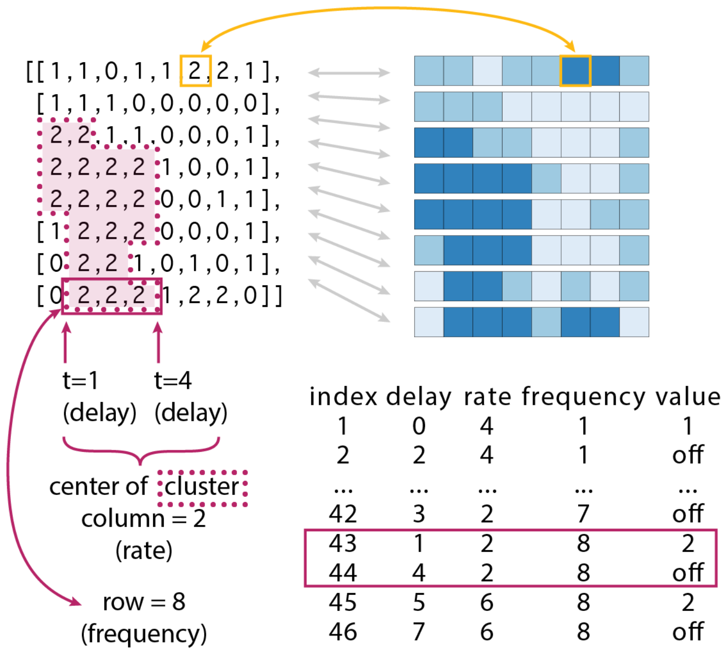

FIGURE 3. Two digital data representations support rendering of the auditory geographic map types. An example data set illustrates the relationships between those formats. A matrix representation of the data (top left) forms the basis for rendering the sequential and augmented-sequential map types. Each symbol in the display, i.e., note in the auditory map or square in the visual analogue (top right), aligns with a cell of the matrix (an example of that alignment is highlighted in yellow). A tabular representation of the data (bottom right) supports rendering of the concurrent map type. Each row of the table reflects a change to the audio-synthesis parameters that is required to encode the underlying data; changes are specified in physical properties of the sound wave: delay (onset and offset time of the sound event), note rate within the sound event, and frequency of the musical note(s).

from the matrix representation one cell at a time and follows a pre-determined order to traverse the entire space (64 grid cells; see geospatial data dimensions, described below). The size of the matrix-based digital data file scales proportionally with the size of the geospatial data set, sensitive to both data resolution and spatial extent (here fixed to an 8×8 grid). In contrast, an alternative data format better supports audio rendering in which multiple dimensions of the audio change at the same time, and multiple audio notes play concurrently. Rather than conforming to the established format for raster geospatial data, a table records audio parameter changes at each time-step of the rendering and improves efficiency for rendering the concurrent map type (notably the transformation between the matrix and tabular formats and its inverse preserve data for the 64 data records). The size of the tabular digital data files grows with the complexity of spatial patterns within the data (as measured by the number of parameter changes need to represent the data, and related to the degree of autocorrelation in the data, Olson, 1975, described in the context of visual complexity). Sampled waveforms in digital files store all rendered auditory geographic maps, regardless of the auditory map type. The file format conforms to the waveform audio file (WAV) format. The rendered files are all the same size, sharing the same sample rate (44,100 samples per second) and duration (56 seconds). The rendered auditory geographic maps are the result of applying three audio parameter mappings to the digital representations of the geospatial data.

Design Constraints

Requirements stemming from the experimental design choices constrain design of the auditory symbology. Limiting the resolution of data across all three dimensions increases experimental control and makes initial evaluation feasible (see Chapter V: Discussion for reflection on the impact on validity). One advantage of the low resolution raster data set is the small number of data points it contains (sixty-four grid cells in an eight-by-eight square). As the duration of sequential auditory geographic map playback scales with the size of the data set, smaller datasets (and thus shorter playback duration) make room for multiple maps in a one-hour session with study participants. At five hundred milliseconds per data value, 64 sequential sound events require at least 32 seconds; temporal overlap in the concurrent map symbology reduces the total duration. A complementary advantage of the raster data format is the ability to break up the data into natural subsets of equal size (rows). The neuroimaging data collection creates substantial acoustic noise which can be restricted to short bursts and interleaved with periods of relative quiet during which to play subsets of the data (or in the case of the concurrent map type, a repetition of the full map). Inserting extra time between playback of contiguous rows, and during which scanner noise occurs, accommodates the unique needs of the neuroimaging methodology with a concomitant increase in the total duration of an auditory map (approximately one minute duration for an instance of the sequential map type with 500 millisecond sound events). With these constraints in mind, the next section describes the adopted audio parameter mapping.

Audio Parameter Mappings

The auditory symbology assigns the three data dimensions to dimensions of the auditory display following a parameter mapping sonification approach. From among the widely recognized dimensions of audio that software parameters can control, the audio symbology uses temporal order, frequency (pitch), note rate (tempo), and amplitude (loudness). Within the organizational framework of the geographic data cube, data and display dimensions are associated with a category of data (space, time, and attribute), and each mapping occurs either within a category or across category boundaries. The two spatial data dimensions are mapped to temporal order, frequency-note rate pairs, or both. The attribute dimension of the data is mapped to amplitude. These geographic data hold time constant,[5] and participants interpret spatial patterns in the generated data without the context of a timeline. This section describes each parameter mapping, and Table 1 and Table 2 summarize the details.

Two distinct approaches map the spatial dimensions of the geospatial data to dimensions of the auditory display. In one approach, spatial location is mapped to time. The sequential and augmented-sequential map types assign the two spatial data dimensions to positions within a linear sequence that plays out over time. A virtual cursor traverses two-dimensional geographic space sonifying each data point that it encounters along its path. This approach to representing spatial dimensions neither uses spatialized audio, nor encodes location of the spatial data in any other attributes of the sound.

[5] With no temporal data to represent, there is no particular benefit to reserving the time dimension of the auditory display for a within-category parameter mapping. However, ceding the time dimension of the auditory display for representation of spatial data dimensions could have a knock-on effect on choices when design guidelines are generalized to more diverse real world maps, eliminating a use of the display's time dimension in a way that is analogous to animation in visual maps. The potential requirement to display time in an auditory geographic map points to a need to understand the representation of spatial data dimensions in audio on spatial or attribute dimensions of the auditory display.

TABLE 1. The three auditory map types use two approaches to parameter mapping sonification to encode spatial data: time-based (sequential and augmented-sequential map types) and attribute-based location (augmented-sequential and concurrent map types). While the underlying data systematically modulates the selected auditory dimensions, all perceptible qualities of the sound are present in an auditory stream; those characteristics of the sound that are explicitly held constant are also listed in the table. The table reports time in milliseconds (ms).

	Cardinal Reference	Audio Dimension	Description	Values
Time-based Location	north-south	temporal order *among* rows	start in the north, progress to the south	0–49000 ms, by 7000 ms (onset relative to map)
	west-east	temporal order *within* row	start in the west, progress to the east	0–3500 ms, by 500 ms (onset relative to row)
	constants:	frequency (pitch)		C_4 (261.63 Hz)
		note duration (rate)		450 ms (1 note per cell)
		temporal envelope	attack, decay	75 ms (0.15×note duration)
			sustain	300 ms (0.6×note duration)
		event duration	total duration of each sound event	500 ms
Attribute-based Location	north-south	frequency (pitch)	high in the north, low in the south	C_3–E_5 (130.81–659.26 Hz)
	west-east	note duration (rate)	slow in the west, fast in the east (value based on regions with homogeneous data)	69–550 ms (1–8 notes per cell)
		temporal order *among* columns	start in the west, progress to the east	0–438 ms (onset relative to cell)
		temporal envelope	attack	22–175 ms (0.35×note duration)
			decay	9–75 ms (0.15×note duration)
			sustain	38–300 ms (0.6×note duration)
	constants:	event duration	total duration of each sound event	500 ms

TABLE 2. Within the parameter mapping sonification, all three auditory map types use amplitude of the sound wave to encode the non-spatial data (population).

Data Value	Audio Dimension	Description	Values
population	amplitude (loudness)	greater amplitude (less attenuation) represents higher population value	[0.10, 0.40, 0.85] (fractional attenuation)

Consistent with an established precedent (e.g., Zhao, B. K. Smith, et al., 2004) a time-based auditory symbology reduces planar location to a linear stream, and one sound event plays for each of the sixty-four raster grid cells following an English reading order (left to right, and top to bottom). Each sound event lasts 500 milliseconds (450 millisecond on time with a smooth attack and decay), resulting in a four second row duration. A quiet pause (three seconds[6]) separates one row from the next, and provides an interval during which the scanner noise occurs without masking the stimulus audio. Relative position *within* a burst of eight sound events (one sound event for each cell in the row), and *among* the eight bursts of sound events (one burst of sound events for each row) indicates spatial location of the data. All of the sound events are rendered sequentially without temporal overlap; the attribute data value determines the amplitude of the sound, and all other parameters are held constant.

The second approach, the concurrent map type, assigns spatial location to a combination of time and attribute.[7] Frequency, which is perceived as pitch,

[6] Map playback starts 500 milliseconds after the one acquisition, and ends 500 milliseconds before the next volume acquisition starts (Perrachione & Ghosh, 2013). Combined with a two-second acquisition time (TA, see below) this necessitates a three second pause between bursts of the auditory stimulus.

[7] The design process makes an effort to identify an attribute-attribute pair of audio display dimensions that could serve as a orthogonal axes of a reference frame in which to represent two spatial dimensions of the data. The effort to identify, implement, and automate combinations of

encodes one spatial dimension: vertical (latitude, or northing). Note rate, which results from subdividing a sound event into an integer number of equal duration notes and gives rise to a perceived quality of slow or fast, encodes the second spatial dimension: horizontal (longitude, or easting). A published finding in the neuroscience literature motivates this particular pair of auditory dimensions. Among macaques, Baumann et al. (2011) found that frequency selective neurons were spatially arranged in a gradient (tonotopy; see also Hall, 2006) that was orthogonal to that of energy over time ("temporal or periodotopic dimension"). Appreciating that evidence of tonotopic and periodotopic arrangement of neurons in non-human primates does not necessarily generalize to humans, if such an arrangement exists it could influence the uptake of spatial information through audio and provides tenuous support[8] for the parameter mapping in the absence of other concrete guidelines for the design of auditory map symbology. Encoding location in frequency-note rate pairs excludes any spatial audio cues.

Independent of the temporal order, frequency encodes location along the north-south axis (vertical, or within columns). The frequencies range from 659.26 Hz (E_5) to 130.81 Hz (C_3). Intermediate rows were encoded in frequencies drawn from an arpeggio of the C-major scale, intervals that fall well above the threshold of noticeable differences (Deatherage, 1972, citing Shower and Biddulph, 1931). High frequencies indicate locations in the north, or the top row of the raster grid; lower frequencies indicate locations toward the south, or the bottom row of

attribute parameters pairs, however, demands a greater time investment in design than is available within this dissertation project.

[8] As observed of retinotopic organization of neurons supporting vision (see footnote 6 of Chapter II: Background, above), the spatial arrangement of neurons in lower level perception do not necessarily predetermine or uniquely predict subsequent higher level processing of sensory input.

the raster grid. Intervals in the major scale also ensure that any combination of notes create neither dissonance (as defined in a Western musical tradition), nor perceptual effect of a beat frequency (Campbell & Greated, 1987), which could confound perception of detection of the note rate (see below).

Also independent of order, but delineated in time, the rate of notes within sound events represent eight discrete locations along a west-east axis (horizontal, or within rows). The note rate varies between one and eight notes per sound event (550 and 69 milliseconds per note, respectively), and note duration includes a smooth attack and decay, which overlap with those of temporally adjacent notes. Note duration within each sound event is inversely proportional to the index of the horizontal location.[9] The auditory symbology assigns slow rates (one note per sound event) to locations in the west, and fast rates (eight notes per sound event) to locations in the east. The minimum change in rate (12.5%) between levels of the note rate is almost twice the threshold for detectability (Drake & Botte, 1993), and the notes are long enough to convey perceptible changes in pitch (Turnbull, 1944). Finally, a single note rate represents contiguous cells that share the same attribute-data value. Grouping adjacent cells that share the same data value (homogeneous cluster) favors the representation of general spatial patterns, which is the objective of the geographic maps in this study, over individual data values. The onset and duration of a sustained sound event encode the east-west extent of the group within each row; the note rate represents the location of the center of the overall group (see Figure 3, above).

[9] In this design process, alternative rhythmic arrangements introduce complexity to the auditory symbol design without providing a substantial benefit. Leveraging such emergent patterns in the design of auditory geographic maps is left to future exploration.

Notably, the name "concurrent" indicates overlap, but does not strictly require simultaneous or instantaneous sound events. The staggered onset of the sounds events occur on 500 millisecond intervals,[10] which shares a similar structure to the 500 millisecond sound events of the sequential and augmented-sequential map types and avoids confusing the staggered onset with the repeated notes of the note rate coding. Introducing staggered onsets requires a strategy for choosing the temporal delay (integer multiple of 500 ms) for sound events. Because the frequency-note rate pairs provide sufficient location information, the staggered onset values could be randomly assigned (akin to jitter used in visualizations to avoid occlusion of overlapping graphic symbols). However, systematic assignment provides an additional parameter of the auditory display that can be controlled as part of the design. Staggered onset values that redundantly encode horizontal position within the geospatial data create a metaphorical sweep or glance across the map from west to east. With the temporal overlap, the concurrent auditory map type encodes the entire map in a single four-second audio burst (equivalent to the duration of a single row in the sequential map type).

Both of these approaches to mapping spatial data dimensions to auditory display dimensions cross conceptual boundaries of the geographic data cube. The sequential and augmented-sequential map types assign spatial data dimensions to a temporal display dimension; the augmented-sequential and concurrent map types assign spatial data dimensions to attribute display dimensions. Although a within-category mapping is possible, since human hearing detects egocentric location cues of a sound source, this auditory symbology does not employ spatial cues. Instead of spatial cues, the chosen auditory symbology gives the map reader control over their

[10] The human auditory system has the ability to discriminate differences in onset time at a much higher resolution

experience through selective attention. Similar to the ways that someone viewing a map directs their gaze to collect fine visual details from spatial locations in a graphic display, someone listening to a map consciously attends to fine auditory details on individual dimensions of a sonic display, or multidimensional patterns that emerge from concurrent representation of display dimensions.

Across all three auditory map types, the attribute dimension of the geospatial data is encoded in an attribute dimension of the audio: amplitude. This mapping stays within the conceptual category of attribute. Amplitude, which is perceived as loudness, encodes classed attribute data values at three discrete levels.[11] Fractional attenuation values between 0.1 and 0.85[12] of a reference value of 90 dB implements the variable amplitude levels; the baseline volume of the speakers or headphones matches participants' subjective preferences on each day of data collection. The middle fractional attenuation value (0.4) is numerically offset toward the low end of the range to produce a subjective perceptual effect of lying halfway between the extreme values. Polarity of the mapping aligns data values with proportional amplitude values, following an analogy with the noise made by groups of people: large groups with many people are (often) louder than small groups. Further perceptual scaling through frequency-based attenuation reduces the impact of interaction between frequency (used in the augmented-sequential and concurrent map types to encode location) and perceived loudness.

---

[11] Notably, struggles to identify specific amplitude values, and that perceived loudness varies considerably between individuals. However, at least in the case of visual geographic maps, participants' ability to tell sequential symbol levels apart has been observed to have little impact on uptake of general spatial patterns (MacEachren, 1982b). The auditory symbology in the three auditory geography map types designed for this study requires listeners to perceive relative differences, but does not necessitate determination of absolute values.

[12] The maximum amplitude value was less than one to avoid a clipping artifact in the digital waveform (amplitude of the waveform exceeds the dynamic range of the digital wave format), which appears when multiple full volume sound events occur simultaneously.

Given the above parameter mappings (and the seven-second acquisition period, TR, of the scan sequence, described below), the auditory symbology for all map types produce geographic map artifacts with 56 second duration. And auditory map design tightly couples with timing of the acquisition, an important consideration for the detection of associated neural activity (Hall et al., 1999). In the sequential map type, presentation of 64 data values (an 8×8 raster grid) fills eight acquisition periods, one for each row. The temporal overlap in the concurrent symbology shortens the time needed to display the entire geospatial data set to four seconds, which fill one acquisition period. Repeating the map eight times, the total duration for concurrent map playback is 56 seconds. Multiple repetitions of the concurrent map type also allow map listeners to selectively attend to different features each time.

Evaluation of Geographic Map Designs

Behavioral and neuroimaging methods evaluate the three auditory geographic map types described above. Justifying the design decisions with theories based in psychoacoustics and harnessing software functionality to implement audio synthesis, however, are just the first step. Understanding the extent to which those designs are effective in their role of communicating spatial pattern is equally important. As increasingly recognized over the last century, scientific methods to investigate map cognition help formalize and systematize the long held belief that map design influences map use, and that the influence extends beyond low-level sensory perception (P. J. Gersmehl & C. A. Gersmehl, 2006; MacEachren, 1995; Montello, 2002): perception is necessary, but alone is not sufficient to accomplish map reading. Borrowing from psychology and neuroscience, this study observes the effect of the use of time in auditory geographic map design through two lenses. Behavioral evaluation indicates where symbolization leads to performance differences. Physiological measures, such as neuroimaging, provide insight into the reasons behind the differences. After grounding the methodological approach to geographic map evaluation in literature, this section describes the test instrument and planned analyses.

Behavioral Testing

Within the overarching objective to improve design of auditory maps, this research seeks to examine the variability in responses to auditory map designs in relation to the temporal arrangement of the audio. A subsequent behavioral action is one part of that response, which is widely used to study spatial cognition and map use (see review in Montello, 2016). Given that effectiveness of communication is a necessary criterion for map evaluation, a behavioral action that relies on information gleaned from the map constitutes an observable metric of the directional relationship between stimulus and perception. Further, such a metric also provides indirect evidence that listeners are attending to the given task, particularly in the neuroimaging setting, which is susceptible to non-task distractions. The behavioral testing component in this study seeks to measure the influence temporal arrangement has on communication effectiveness.

Behavioral methods provide metrics through which to evaluate (auditory) maps by measuring effects on communication of spatial distribution. While Muehrcke (1973) notes that "in theory, it should be possible to measure the communication efficiency [...] by simply measuring the correlation between input and output information" (p. 190), in practice quantifying inputs and outputs is not simple. A measure of correlation requires a specific task context and depends on quantifiable and comparable values across both the *input* map and resulting *output* mental representation. Quantifying those inputs and outputs can be problematic. Even through quantitative data fill the raster grid, the target or critical characteristic(s) of that data set may be complex or ambiguous. As input to a geographic map display, aggregate characteristics at the level of the data set may be of greater interest than individual data values. In this study, the input is a raster data set that manifests spatial patterns; Appendix D: Experimental Stimuli lists quantitative metrics that characterize those patterns. Measuring the resulting mental representation, too, is difficult. Following one component of a task described by MacEachren (1982b), a paired comparison served as a way to externalize or express one feature of a mental representation. Participants make a judgement about the relative magnitudes of data values at pairs of locations under two memory conditions: recall and active listening. Successful query of a mental representation suggests that a geographic map display has communicated sufficient task-critical information. The "correlation" between the input and the output in this scenario then becomes a record of whether or not the response to the paired comparison matches the correct directional inequality according to an objective comparison of input data values. Beyond a measure of accuracy, the behavioral data also reflect whether or not the participants attend to the task even in the confined and noisy setting of the MR scanning environment.

The behavioral testing approach to evaluating map effectiveness makes two key assumptions. First, it assumes that the only way to successfully answer the task question is by extracting data from the auditory display and forming an accurate mental representation of that information. With a paired comparison task, there is a high likelihood of selecting a correct answer by guessing. To mitigate the negative impact of this assumption, the analysis compares performance across conditions with the expectation that any difference reflects both differences between the conditions and differences in guessing that are modulated by the conditions. Second, there is an implicit assumption in many quantitative analyses that all people perform the task similarly and comparison against a group average value is meaningful. There is, however, a high degree of variability between individuals. To mitigate the influence of that variability, a within-subject design sought to measure relative differences in effectiveness.

While the data from these behavioral methods provide insight into map effectiveness, they cannot indicate the mechanistic causes of any observed differences. And literature from human computer interaction has documented people's inability to both consciously recognize and subsequently articulate *why* a display works (Nielsen, Clemmensen, & Yssing, 2002). Neuroimaging, the second prong of this research approach, complements the behavioral data by providing an opportunity to investigate (unconscious) responses by observing underlying neural activity that occurs during map listening.

Neuroimaging

Neuroimaging offers a quantitative metric for empirical evaluation of auditory map effectiveness by targeting brain responses to geographic map listening. Neuroimaging data reflect a quantitative measure of blood flow through the blood-oxygen level dependent contrast (BOLD), which is an indicator[13] of localized

---

[13] Despite inherent uncertainty associated with using BOLD as a proxy for neural activity, fMRI is "currently the dominant paradigm for assessing behavior-related brain physiological changes in humans" (Ekstrom, 2010, p. 234). There is a strong reproducible correlation between

brain activity (Ogawa et al., 1990). Data collected through neuroimaging provide a unique lens through which to observe participant reactions to map listening.[14] While listeners' conscious impressions of the display relate to perceived and physical map characteristics, people often struggle to both identify and describe usability issues in software interfaces (among both participants and researchers, Nielsen et al., 2002; Walker and Kramer, 2006). And, unlike the assessment of visual maps, which can be investigated by eye tracking (see Montello, 2002), the study of auditory maps and underlying attentional control, which is served by covert attention shifts, cannot be directly observed. Understanding how the auditory stimulus modulates the neural response to the temporal arrangement support the expanded use of audio in map design.

Neuroimaging data measures BOLD response as an indicator of the neural response to the auditory geographic maps. Two pairs of contrasts[15] evaluate differences in neural responses to listening to the three auditory map types. Two contrasts (A>B and A<B) measure differences (activation and deactivation) in neural response attributable to the level of explicit location information: sequential compared to augmented-sequential. The other two (C>D and C<D) measure the

---

the BOLD signal and neural activity as measured by single cell recordings (e.g., in the visual cortex Kanwisher, 2010). (For a detailed history and review of fMRI, see Amaro, Jr. and Barker, 2006; Bandettini, 2012; Cohen and Schmitt, 2012; Maus and van Breukelen, 2013.)

[14] Cartographers and geographers have established a precedent for using fMRI to explore the uptake of information from a cartographic map (e.g., Lobben et al., 2009; Rozovskaya & Pechenkova, 2012; Zhang, Copara, & Ekstrom, 2012) and to understand spatial thinking (e.g., Auger, Mullally, & Maguire, 2012). Studies using fMRI have the potential to identify mechanisms behind observed behavioral differences on geospatial tasks (Lobben, Lawrence, & Pickett, 2014, 2005).

[15] The experimental design supports an additional analysis to assess the factorial interaction (Amaro, Jr. & Barker, 2006) between auditory geographic map type and response condition was considered, but the collected data is insufficient to support such analysis (short duration of response activity, often a single TR; limited repetitions, only one instance of each condition of the 3×2 factorial design within each run). The influence of auditory symbology relative to response condition is left to future work.

difference in response attributable to temporal arrangement: augmented-sequential compared to concurrent. All fMRI measurements are relative (see discussion in Aguirre & D'Esposito, 2000), and comparison between paired experimental conditions removes the need for an arbitrary control condition or rest. The two step difference between the "established" sequential audio maps and the novel concurrent audio maps supports isolation of the influence of temporal arrangement from the influence of different levels of information. These four contrasts identify whole brain activation patterns.

Design of the neuroimaging study relies on three important assumptions. First, the use of contrasts that represent subtraction of neural activity levels assumes that processing is supported by separable cognitive actions, or pure insertion (i.e., subtraction of two processes produces a measure of a cognitive task, Donders, 1868/1969; and see Aguirre and D'Esposito, 2000). Even through it is expected to be invalid, particularly in the presence of scanner noise (Hall et al., 1999) or when cognitive resources are overwhelmed, violation of the pure insertion assumption does not preclude informative results (see Amaro, Jr. & Barker, 2006) To minimize the impact of confounds stemming from interrelated and non-separable brain activity, a third map type (augmented-sequential) holds constant as many aspects of the display as possible while retaining the ability to compare with established map designs. Second, the selected general linear model (GLM) analysis approach assumes a linear relationship between stimuli and the BOLD response (Mumford, 2010, and see neuroimaging analysis described below). Physiological responses such as adaptation contribute to a non-linear relationship between the stimulus and neural response. Documented violations of the linearity assumption, however, are not sufficiently problematic to reject the GLM approach

(Boynton, Engel, Glover, and Heeger, 1996, observed in visual cortex). Third, the analyses rely on functional specialization in the brain. Neurons that are selectively active in response to specific stimuli or cognitive tasks tend to be co-located in the brain. As an ongoing topic of debate, contemporary fMRI techniques provide evidence that continues to support functional specificity, although to a more conservative degree (Kanwisher, 2010). Finally, the group level analysis relies on consistency in spatial layout of functional areas in the brain across individual participants. Although general patterns of functional activation are comparable across individuals, functional activation does not strictly align with anatomical regions and there may be strong individual differences.

Taking the three audio map types and a methodological approach that pairs behavioral and neuroimaging methods, a empirical study evaluates participant responses to a map listening task.

Experimental Design

This section describes the implementation of a test instrument that uses behavioral and neuroimaging methods. A block design captures information about sustained neural activity related to processing map content (in contrast to the lower-level process of sensing audio, which relates more directly to the onset or change of stimuli). The block design detects sustained neural activity better than event related design (e.g., Visscher et al., 2003, evaluated in the context of sustained attention to changing visual stimuli). Encoding and rehearsing spatial patterns that the auditory display presents requires sustained activity beyond simple detection in the sensory cortex. Regular presentation order in the block design benefits the detection of activation which can be suitable for initial stages of an investigation (e.g., identification of regions of interest for further investigation Aguirre and D'Esposito, 2000; although such regularity is problematic in studies that seek to model the shape of the hemodynamic response function, see Liu, Frank, Wong, and Buxton, 2001).

The study follows a protocol that was reviewed and approved by the University of Oregon Institutional Review Board.

Target Population

The target population is adults with normal hearing. Recruiting targets a sample population from a broad cross section of ages (18-65 years old) and experience (no limiting selection criteria; see Table 3). Although accessibility for people who are blind is a potential application domain, this study uses a combination of auditory and visual materials, which precludes generalization of the results to a general population with, for example, more diverse levels of vision. The study recruits participants around the University of Oregon campus, but has no requirement that participants be students or have any particular educational background or training.

The target sample size is twenty four of participants. The documented lack of metrics to serve as estimates in power analysis (Q. Guo et al., 2014) coupled with the uniqueness of the selected scan sequence and the novelty of the task precluded an a priori power calculation. Instead, the selected target sample size balances general guidelines with typical study sizes reported in the literature.

TABLE 3. Eligibility criteria for participation in the study aim to accept a broad range of people, while taking into account the task design and safety considerations for the MR scanning environment.

Criterion	Justification
18–65 years old	Restrict the study population to adults, who can autonomously give informed concent
normal hearing	Stimuli (auditory maps) were presented aurally
normal or corrected to normal vision	Instructions and response options were presented visually
no permanent metal devices or implants, and not currently pregnant	Safety and precautionary considerations for the MR scanning environment

Procedures

Each participant attends two sessions on separate days. At the start of each session, participants provide their informed consent to participate, confirm that they met the eligibility criteria, and (in the first session) confirm their availability for both study sessions. As part of the informed consent process, participants indicate whether or not they consented to the public release of their de-identified study data. Participants receive compensation for their time at a rate of $25 per hour for a training session, and $35 per hour for a scanning session.

The first session, a training session, occurs in a private computer lab setting and lasts up to an hour. To start the session, the researcher provides information about the study and participants answer demographic questions (see Appendix A: Demographic Questions). The researcher then introduces the three audio types and the map interpretation task, encouraging participants to develop a general idea of any spatial patterns that the data create. After listening to the auditory geographic map examples, participants verbally describe what they have heard. After hearing descriptions and examples of the three audio types, participants complete a practice run consisting of a set of six maps (trials), one for each condition. Participants are encouraged to ask clarifying questions as they arise. Throughout the session, participants provide unstructured feedback on the overall design of the three audio types.

On a later date, participants complete a scanning session in the Lewis Center for Neuroimaging on the University of Oregon campus. At the start of the session, and after again indicating their informed consent to participate, participants listen to samples of each of the three audio types. Participants then review a safety questionnaire with the MR Technologist, adjust a pair of MR safe glasses (as needed for corrected to normal vision), adjust the volume for stimulus presentation through MR safe headphones to a comfortable level, and are helped into the scanner.

The scanning session comprises two functional scans and four reference scans (Figure 4). Functional images record the BOLD signal while participants listen to and make judgements about the maps. During each functional scan, participants complete one set (run) of six questions (trials). The reference scans, which provide information to assist with image processing and interpretation, are acquired before, between, and after the functional scans. The first acquisition, a short reference scan, records participant position within the scanner and head coil, and serves as participants' first exposure to the sounds of the MR scanner. Between the two functional runs, an anatomical image offers an opportunity to rest between task runs, and the resulting data facilitates registration of the functional images to standard space for the group analysis. Finally, a pair of field strength images capture information to create a field map that is used in the pre-processing step of analysis to mitigate inhomogeneity in the main magnetic field. The researcher checks in with the participants between scans, speaking through the headphones and listening for participants' verbal responses through the intercom system. The overall scanning session lasts approximately one hour.

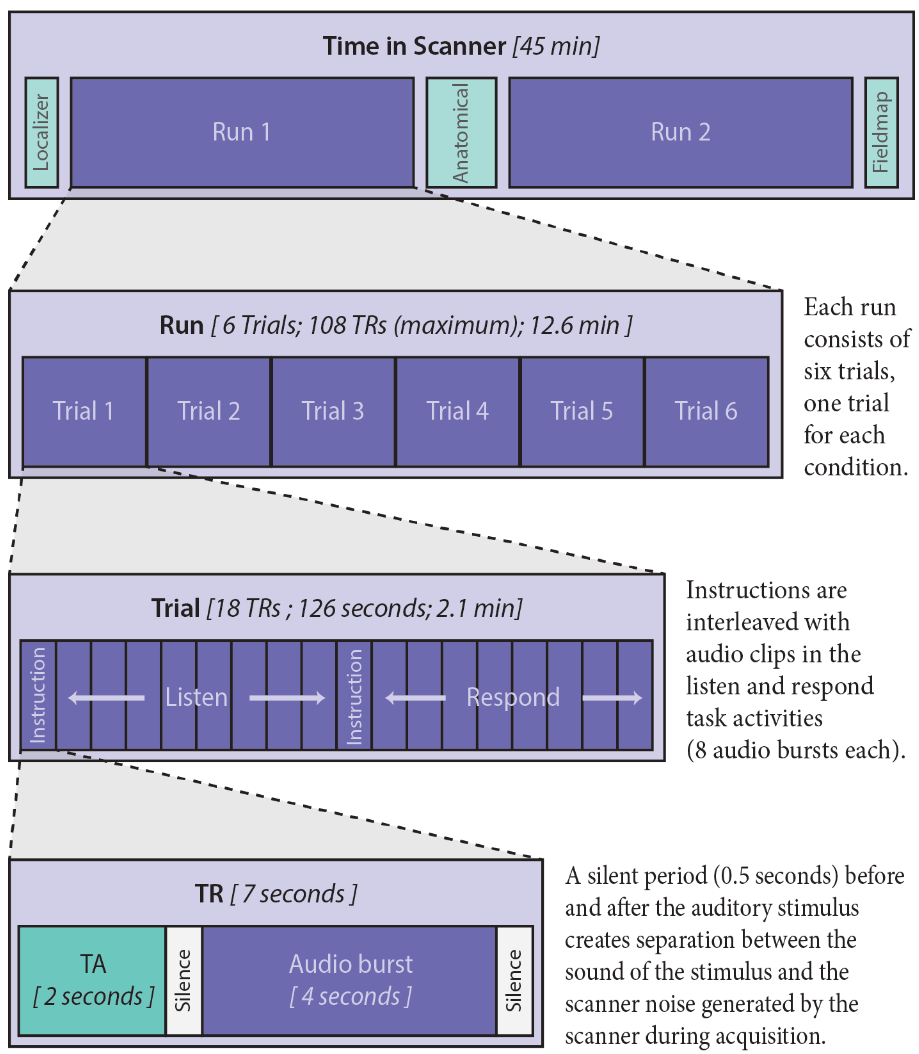

FIGURE 4. The structure of the scanning session builds on the sparse sampling TR of seven seconds. Each task activity – listen and response – spans eight TRs. Paired task activities make up trials, and six trials comprise a functional run. Three reference scans and two functional runs lead to approximately 45 minutes in the scanner.

Test Instrument

The test instrument borrows several elements of the study design described by MacEachren (1982b, hereafter the "original task"). Modifications to the design accommodate the auditory stimuli, adapt to the fMRI methodology, and constrain the scope of the study. The auditory stimuli take longer to display than visual stimuli, limiting the number of maps that can be displayed during a one-hour session. The value comparison task was simplified from the paired comparisons for all possible pairs of regions to a single comparison between two locations.[16] Response options correspond to two buttons in a binary forced choice task. While the extraction of exact data values is one function of map displays, the scope of this study is limited to only "general" map features, which are particularly challenging to interpret from auditory displays (e.g., Alty & Rigas, 2005; Delogu et al., 2010). In contrast to the original study, geographic map complexity, which is measured in this case as the number of levels in the classed attribute data and the degree of spatial clustering, is held constant across map instances.

Recruiting A paper flyer posted around the University of Oregon campus solicits participation. The flyer introduces the topic of the study, states the eligibility criteria (Table 3, above), and directs anyone interested in participating to a web-based form for more information. The researcher contacts respondents via their indicated preferred contact method (email or phone) to provide additional information about the study and to schedule data collection sessions, making an

[16] Although the use of a visual prompt introduces a confound in the form of translation between visual and auditory representations of location, the graphics constituted a robust and familiar way to present location to participants who were sighted.

effort to schedule scanning sessions within three days of that participant's training session.

Behavioral Task A behavioral task measures the degree to which the auditory maps effectively communicate general map information. The training session introduces the behavioral task and testing procedure that is followed during the scanning session. A verbal description of the geographic maps indicate that they represent population, which provides a geospatial context for the simple raster data. The training used terminology such as "the number of people" and "where were there more people" when instructing participants how to interpret the auditory display.

The task consists of listening to a map and then making a judgement about the relative magnitudes of the data values (loudness) at two target locations. In a randomized order, participants listen to two maps in each of the audio map types. A fixation icon appears 500 milliseconds before the start of the auditory map playback in the center of the screen on a white background. The icon remains on the screen during the map playback. The shape of the fixation icon (Figure 5) indicates the temporal arrangement of the upcoming auditory map type; the display gives no indication of the pending response condition. During map listening, participants are free to have their eyes open or closed. After listening to a map, the software displays a reminder of the task instructions and then visually

FIGURE 5. The screen displays a visual fixation icon 500 milliseconds prior to the start of the map listening activity, and throughout that activity. The shape of the icon indicates the temporal arrangement of the upcoming auditory map type: sequential (left) or concurrent (right).

presents two target locations. Similar to the original task, the training primes participants to attend to general map information, or patterns that the data create. The training session introduces the instructions verbally and during the scanning session a graphic reminds participants of the task instruction (Figure 6). If the participants close their eyes during map listening, presentation of the instruction image without an accompanying stimulus sound serves as a cue that the map playback has ended. During the response portion of the task the software renders two squares, which correspond with two locations in the geospatial data, on an otherwise blank light colored background (Figure 7). Participants choose the location at which there is a higher (louder) data value, entering their answers using two arrow keys on a keyboard (training session) or button box (scanning session). The target location nearer the west (left) edge of the geographic map is associated with the response button under the index finger of the right hand[17]; the index finger is on the left side of the hand when extended, palm down. Similarly, the target location further to the east (right) was associated with the response button under the middle finger. Responses were made under two memory conditions. After one of the two instances of each audio type, the map stimulus was silent during the response activity (*recall*; Figure 8); in the other instance, the map data played a second time (*active listening*). The effective display area on the screen for all graphical displays (fixation icon, instructions, and response prompt) has a square aspect ratio.

[17] A left handed version of the response buttons establishes a similar egocentric association between the position of the response options within the geographic map and the finger assigned to each response option.

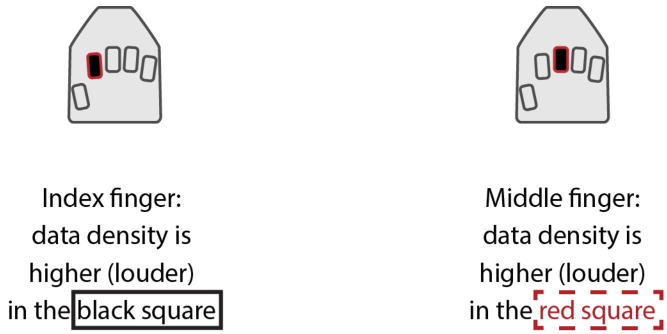

FIGURE 6. A graphic reminder of the task instruction precedes each response activity. The graphic depicts the button box hardware (or arrow keys in the computer-based training session, not pictured), and visually highlights the mapping between physical buttons and the response options. This instruction graphic is specific to the right hand, and a version of the graphic depicting the left hand is also available.

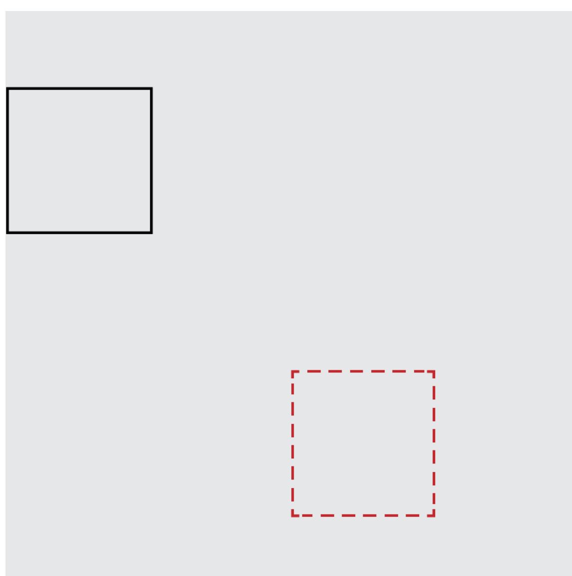

FIGURE 7. The visual display of two target locations indicates the start of the response activity. This example graphic shows a location on the upper left edge in black and a location on the lower right in red. The black and red squares corresponded to the index- and middle-finger response buttons, respectively (right hand).

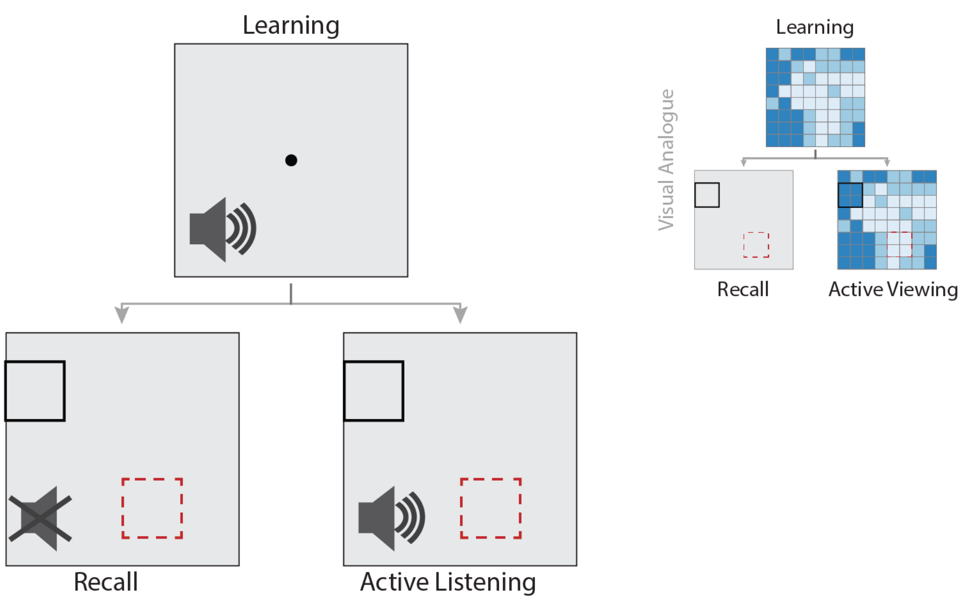

FIGURE 8. The behavioral task consists of two parts. First, participants listen to an auditory map rendered in one of the three map types. Second, a visual prompt indicates two locations within the map that participants are asked to compare under two memory conditions. A visual analogue of the task (top right) provides further illustration, but is not available to participants during the task.

Scan Sequence A custom sparse sampling scan sequence addresses the unique needs of presenting auditory stimuli in the MR scanning environment (Belin, Zatorre, Hoge, Evans, & Pike, 1999; De Martino, Moerel, Ugurbil, Formisano, & Yacoub, 2015). This custom scan sequence balances the timing needs with constraints of the data collection technique. A Siemens Skyra 3T scanner in the Lewis Center for Neuroimaging on the University of Oregon campus collected neuroimaging data. An MPRAGE sequence (TR 2500 ms, TE 3.43 ms, flip angle 7°, voxel size 1 mm $\times$ 1 mm $\times$ 1 mm, matrix size 256 $\times$ 256, 176 slices) captures anatomical images. And a spin-echo echo planar imaging (EPI) sequence (epif2d1 64, TR = 7000 ms, TA = 2000 ms, TE = 27 ms, flip angle 90°, slice thickness 4 mm, voxel size 3.25 mm $\times$ 3.25 mm, matrix size 64 $\times$ 64, 33 slices) captures functional images. Although acceleration techniques promise improved spatial resolution, exploration of one such technique (multi-band acquisition) reveals an interaction with the sparse sampling that precludes its use (see Appendix B: Scan Sequence Development: Multi-band Acceleration).

Using this configuration, two instances of the EPI sequence acquire data for two functional runs. Each run consists of six trials, which in turn break down into listen and respond activities (Figure 4, above). The listen task produces 48 volumes for each run, eight volumes for each condition. The number of volumes collected during the response task varies depending on how quickly a participant responds. The maximum duration of a run is 12.6 minutes, which is within the typical duration of a functional scan (five to 15 minutes, Maus & van Breukelen, 2013).

During image acquisition, vibration in the gradient coils and other scanner hardware create substantial noise, which can exceed 100 dBA (see reviews by

McJury & Shellock, 2000; Moelker & Pattynama, 2003). This noise poses two potential problems. The first consideration is the impact of acoustic noise from the scanner on participants. Even when the noise levels fall below standard thresholds for causing permanent hearing damage during a single scanning session (McJury & Shellock, 2000; Shellock & Crues, 2004), scanner noise can still cause discomfort. Second, the acoustic scanner noise interferes with perception of auditory stimuli. Scanner noise has been found to increase sensitivity thresholds for the perception of pure tones (Ulmer et al., 1998), decrease magnitude of the neural response to pure tones (including interaction with relative timing of scanner noise and stimulus presentation, Langers, Van Dijk, & Backes, 2005), and distract attention away from the task (Moelker & Pattynama, 2003) with implications in working memory (Tomasi, Caparelli, Chang, & Ernst, 2005, letter recognition task). This study uses passive hearing protection to mitigate the high level of acoustic noise.[18]

Magnetic resonance (MR) physics offer some intuition about what these choices in scan sequence design mean to data collection and the resulting data set. Functional magnetic resonance imaging (fMRI) acquires digital images that create a three-dimensional representation of the brain (or body). The scanner cannot acquire digital images instantaneously, and instead records a series of two-dimensional slices, which are subsequently reconstructed into a three-dimensional volume. The number of slices depends on the spatial extent of the volume and the thickness of the slices: greater spatial extent necessitates either a greater number of slices or thicker slices with lower spatial resolution. To encode slice location in the MR signal, gradients systematically distort the main magnetic field and radio

[18] Alternate approaches to reducing scanner noise exist, ranging from passive hearing protection to modulation of sequence parameters (see reviews by McJury & Shellock, 2000; Moelker & Pattynama, 2003) and design of scanner hardware (e.g., Bowtell & Peters, 1999). But these approaches are beyond the scope of this study.

frequency (RF) pulses selectively excite molecules within a transverse plane (xy-plane). Intensities at locations within the two-dimensional slice are reconstructed from measurements in k-space (angle, phase, and amplitude of spatial frequencies). The number of samples recorded from k-space relates directly to the in-plane spatial resolution of the resulting image, and the time required to acquire a single slice. The number of slices and the time to acquire a slice combine to determine the time needed to collect a single volume (acquisition time, TA). The time interval between volume acquisitions (repetition time, TR) depends on the experimental design. While scanner hardware and laws of physics dictate a minimum TA, TR is a flexible scan sequence parameter.

Sparse sampling inserts a delay between the end of one volume acquisition and the start of the next. The schedule on which the scanner acquires volumes (TR) is slower than strictly necessary to acquire the image (TA). But such a delay creates an effective quiet interval (although notably not silent, Moelker and Pattynama, 2003). Further, reducing the baseline level of activity in the primary auditory cortex provides a higher signal to noise ratio (Hall et al., 1999), but it also reduces the number of volumes that are acquired in the fixed amount of time. In this study, a TR of seven seconds was selected. This slow TR (within the range that Perrachione and Ghosh, 2013 recommend) produces quiet intervals of time that are free from the noise of the gradient coils, and during which to present auditory stimuli. Notably, however, the hemodynamic response peaks several seconds after the onset of a related stimulus (Belin et al., 1999; Talavage, Gonzalez-Castillo, & Scott, 2014). The neural activity of interest likely flows across multiple acquisitions and is contaminated by the sensory response to acoustic scanner noise

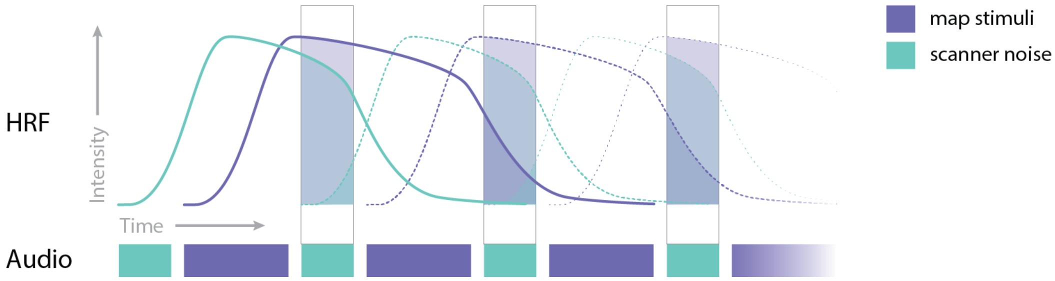

FIGURE 9. The peak of the hemodynamic response, illustrated as an approximate hemodynamic response function (HRF), lags behind the onset of the stimulus. In the selected scan sequence and with the three auditory map types, the measured BOLD intensity reflects response to both the stimuli, but residual and new scanner noise also contaminate the measurement.

(Figure 9). A different TR may have captured the BOLD signal nearer its peak (see Hall et al., 1999), but would also extend the playback time for each map.

Sparse sampling inserts extended spin relaxation time between acquisitions The extended relaxation time produces higher intensities in the measured BOLD signal in the first slices of a volume compared to those acquired later in the sequence (see also Appendix B: Scan Sequence Development: Multi-band Acceleration). Although higher intensity at the start of the volume provides higher signal to noise ratio, it also creates an intensity gradient from the superior (top of the head) to inferior (toward the neck) slices in the image, following the slice acquisition order. This intensity gradient confounds direct comparison of intensity magnitudes across superior and inferior regions of the brain,[19] but does not preclude comparison between task conditions within brain regions.

Presentation Hardware In the scanning session, a desktop computer (Mac OSX 10.10.5, 2.7 GHz processor) runs the presentation software. MR-safe

[19] In addition to the spatial gradient of intensities, other spatial inhomogeneities, e.g., spatially heterogeneous shapes of the hemodynamic response function, also challenge the validity of direct comparisons across different brain regions.

headphones,[20] which attenuate scanner noise (passive hearing protection), deliver auditory stimuli. Both ears receive identical copies of the single channel audio. The headphones fit snuggly in the 20-channel head coil. A foam wedge offers additional attenuation of scanner noise as propagated through bone conduction (see Moelker & Pattynama, 2003) and additional foam sponges provide support to help the participants hold their heads still. Participants view a projected screen image through a mirror mounted on the head coil. Participants enter their answers to the behavioral task via two buttons on a button box.[21] Between scans, the researcher communicates with participants using a microphone connected to the headphones and hears participants' verbal responses over the intercom system.

A laptop computer (Mac OSX 10.11.6, 2.53 GHz processor) runs the presentation software during the training session. The internal speakers and built-in screen display the auditory geographic maps and graphic prompts. Participants press arrow keys on the built-in keyboard to indicate their responses to task questions.

Presentation Software Custom software uses PsychoPy (Peirce, 2007, 2009) to automate stimulus presentation. The PsychoPy toolbox provides functionality to present both visual and auditory stimuli and offers microsecond level control over timing. In addition to stimulus delivery, the software records behavioral responses and synchronization messages from the scanner.

The software requires three parameters at runtime to configure the presentation for each session. A unique participant identifier labels each of the

[20] NordicNeuroLab, AudioSystem, http://www.nordicneurolab.com/products/AudioSystem.html

[21] Psychology Software Tools, Inc., Celeritas Fiber Optic Response System, https://pstnet.com/products/celeritas/

FIGURE 10. PsychoPy-based software automates presentation of auditory map stimuli (Peirce, 2007, 2009). The implementation loops though six map instances (trials) during which which participants listen to an auditory map and then make a judgement about the map contents. A trigger pulse synchronizes stimulus presentation with the fMRI acquisition. The flow of Run2 was identical to that of Run1.

output files and serves as a link across data types (demographic, behavioral, and neuroimaging data; see also Appendix C: Data Dictionaries). The type of session (training or scanning) determines the conditional display of the instruction graphic (response via keyboard arrow keys or buttons on the button box). The participants' self-reported preferred hand facilitates interpretation of the coded response events from the keyboard or button box, and determines the conditional display of the response graphic.

The software presents sets of stimuli in three runs: one run during the training session and two runs during the scanning session. A loop controls presentation of the map trials (Figure 10); PsychoPy built-in functionality randomizes the presentation order of the trials. Within each run (training, Run1, Run2), a pre-determined set of six maps were displayed, each map is presented exactly once, and each condition (map type × response) occurs exactly once. The randomized order mitigates learning effects and task transition effects (Maus & van Breukelen, 2013). The three sets of stimuli contain comparable collections of map properties (see stimuli, described below). The order of the two sets of maps used in the scanning session is counterbalanced across participants.

Stimuli Synthesized map data provide experimental control. A custom R script systematically generates map data as point patterns (spatstat::rMatClust; κ=70, scale=0.2 in a square window with side length 1, and μ=20). Counts of the number of points that fall in each grid cell (rgeos::gIntersects) transforms a point patterns into an 8×8 raster grid. Quantile breaks (classInt::classIntervals) class the raster data into three levels. The use of quantiles ensures that each map has approximately equal numbers of cells that belong to each data level (MacEachren, 1982a). Each trial depicts a unique data instance to avoid the possibility that participants recognize repeated data, and characteristics of the data sets are balanced across runs to the extent possible (see also Appendix D: Experimental Stimuli).

Criteria for experimental control include only those instances that possess suitable task response options. A pair of square regions that each cover an area equal to two rows tall and two columns wide serve as target locations for the comparison task (see Figure 7, above). The two target locations are non-adjacent and non-overlapping across both rows and columns. The general direction (northeast vs. southwest) and distance between the two target locations is balanced within each run. To avoid ambiguous comparisons and act as a control for the level of difficulty across instances of the task, each target region contains strictly homogeneous data values, and the data value was “low” in one region and “high” in the other. The relative position of the correct response (western or eastern; corresponding to the higher value within the pair of target locations) is balanced within each run. The script generates data until a sufficient number of data sets satisfy the experimental controls.

A custom Python script uses the Pyo[22] module to render the digital data in auditory geographic maps. The script renders audio offline to ensure consistent stimuli across study sessions. Manual editing (Audacity) removes any instances of clipping. Supplemental audio files Supplementary Files, Audio 1 (sequential), Supplementary Files, Audio 2 (augmented-sequential), and Supplementary Files, Audio 3 (concurrent) provide examples of the auditory geographic map types. The three supplemental files represent a single data set in each of the three auditory-map types; Figure 3, above, illustrates the data that these auditory-map examples represent.

Data

This study produces data in the form of processing scripts and empirical observations. A collection of scripts automate the preparation of stimuli, present stimuli, and analyze empirical data. The scripts that create the auditory map stimuli serve as artifacts to document design decisions. The PsychoPy experiment script details stimulus and task timing. And the R and bash scripts that perform and automate statistical analyses create a reproducible analysis pipeline and document parameters of the analyses.

Empirical data include both behavioral performance and neuroimaging metrics. Behavioral data from both the training and scanning sessions reflect details of the map instances, their presentation order, response values, and response time. Notably, data from the training session is contaminated with substantial noise as participants learned the task and instances of the task were interrupted

[22] Olivier Blanger, Pyo, http://ajaxsoundstudio.com/software/pyo/.

by questions for clarification.[23] The scanning sessions produce collections of image files for each participant. Automated processing converts the images After minimal processing (see neuroimaging analysis, described below), the neuroimaging data and associated metadata conform to the Brain Imaging Data Structure (BIDS, v1.0.2; Gorgolewski et al., 2016) standard. In accordance with participant consent to the release of their respective data sets the data are publicly available through OpenNeuro (https://openneuro.org/datasets/ds001415).

Analysis

The analysis handles the two types of empirical data, behavioral data and neuroimaging, separately. Behavioral data collected during the scanning session addresses the first research question (RQ1: influence of temporal aspect of audio on communication of spatial patterns). Neuroimaging data addresses the second research question (RQ2: contrast in neural activity in response to sequential vs. continuous auditory stimuli).

Behavioral Analysis Analysis examines the behavioral data from the scanning session for evidence of differences in performance between the three different map types. Response time and accuracy provide a quantitative measure of performance under each of the two response conditions. Response time data reflects the time that elapses between the onset of the response prompt (and audio playback in

[23] The behavioral data collected during the training session are incomplete. Multiple trials in the training session were interrupted by questions (five trials) or incorrect advance by the software (one trial); no more than one interruption occurred in a single session. Data associated with interrupted trials reflect not only task performance, as measured by response time and accuracy, but also the time spent asking questions and confusion about the task itself (beyond the target relative judgement of data values). And, due noise in the behavioral data set from the training session, which already consists of a small number of observations, an originally planned comparison between performance in a computer-lab setting and performance in the scanner is not feasible.

the active listening condition) and the subsequent button press. A binary score (correct or incorrect) measures accuracy. Each of the 22 participants contribute two response time and two accuracy values for each of the six experimental conditions: three map types and two response conditions. The analysis conducts separate comparisons for the two response conditions. A Friedman rank sum test (Conover, 1971; Sheskin, 1997; as implemented in R stats::friedman.test), a non-parametric test that accommodates heterogeneous variance and lack of normality in the data, tests the response time data. Four tests assess the degree of difference in response time between paired experimental conditions. Within each response condition, two tests measure differences between concurrent and augmented-sequential map types, and between augmented-sequential and sequential map types were measured. A McNemar test for significant changes (Conover, 1971; Sheskin, 1997; as implemented in R stats::mcnemar.test), a non-parametric test for paired measures with binary outcomes, quantifies the likelihood of differences in response accuracy. This analysis consists of eight tests, which correspond with four conditions (two map types and two response conditions) for each of the two runs (repeated measures).

Neuroimaging Analysis Neuroimaging data provide a quantitative measure of differences in the BOLD response in association with the presentation of three types of auditory geographic maps. The analysis covers standardization of the data format, preprocessing, and statistical analysis.

The first step of the analysis standardizes the format of the data files. MRIConvert (mcverter, 2.1.0 build 440, Lewis Center for Neuroimaging, University of Oregon) converts raw data from the scanner (DICOM[24]) into a standard image format (NifTI[25]) that analysis software accept as input. Utilities from the FMRIB[26] Software Library (FSL, v5.0.10, Jenkinson, Beckmann, Behrens, Woolrich, and Smith, 2012), a software suite for neuroimaging analysis, apply rotations in increments of 90 degrees to orient the images with the standard axes and updates the NifTI headers (reorient2std, FSL utilities) and trim the image to focus on the brain (robustfov, FSL utilities). An automated tool removes personally identifiable facial features from the images (mri_deface, v1.22, FreeSurver, Harvard; Bischoff-Grethe et al., 2007). And JSON-formatted sidecar files accumulate metadata (dcm2niix, v1.0.20171215, Chris Rorden; custom script to update locally defined fields). The BIDS Validator (bids-validator, v0.26.14, International Neuroinformatics Coordinating Facility, Stanford) checks compliance with the BIDS standard (Gorgolewski et al., 2016).

Preprocessing is consistent with recommendations from the literature and leverages functionality from FSL. A custom script extracts stimulus timing information from the log files, creating a tab separated file (consistent with the BIDS format) and a space separated explanatory variable files (EV, consistent with the FSL three-column format). Event timestamps reflect both onset relative to the first scanner synchronization message and an adjusted onset, which tools in FSL expect, that accommodates the sparse sampling sequence by adjusting the midpoint of the TA to align with the midpoint of the TR. An automated processing

[24] Digital Imaging and Communications in Medicine, ISO 12052, https://www.dicomstandard.org/

[25] Neuroimaging Informatics Technology Initiative, National Institutes of Health, https://nifti.nimh.nih.gov/nifti-1

[26] Wellcome Centre for Integrative Neuroimaging (FMRIB), University of Oxford, https://www.ndcn.ox.ac.uk/divisions/fmrib

pipeline (fsl_anat, *beta*, FSL utilities) applies bias correction to the anatomical images and restricts the field of view. Double-gradient echo images yield a fieldmap (fsl_prepare_fieldmap, FSL FUGUE utilities) that represents inhomogeneity in the main magnetic field. The brain extraction tool (BET, FSL, S. M. Smith, 2002) isolates the portion of the each image that represents the brain from non-brain data. FMRI Expert Analysis Tool (FEAT, v6.00, FSL; M. W. Woolrich, Ripley, Brady, and Smith, 2001; M. Woolrich, Behrens, Beckmann, Jenkinson, and Smith, 2004), performs image registration (FLIRT, FSL, Jenkinson, Bannister, Brady, and Smith, 2002; seven degrees of freedom and boundary based registration, BBR), unwarping (FUGUE, FSL, Jenkinson, 2002, 2004), motion correction (MCFLIRT, FSL, Jenkinson et al., 2002), spatial smoothing[27] with a Gaussian kernel (SUSAN, S. M. Smith and Brady, 1997; five millimeter full width half maximum), and pre-whitening (FILM, M. W. Woolrich et al., 2001) to address temporal autocorrelation Visualization tools (FSLView, v3.2.0, FSL utilities; FSLeyes, v0.22.6, McCarthy, 2018) facilitate inspection of the output from each step of the preprocessing.

Two parallel analyses measure statistical significance of directional differences in the BOLD response during the listening portion of the task. One analysis evaluates the influence that the level of explicit location information has on the neural response with contrasts between the sequential and augmented-sequential map types. The other assesses differences attributable to the temporal arrangement with contrasts between the augmented-sequential and concurrent map types. Both analyses use a general linear model (GLM; see Beckmann, Jenkinson, and Smith, 2003) approach to whole brain analysis in three passes.

[27] Spatial smoothing improves the power to detect activity in a limited number of volumes, but does so at the expense of spatial resolution (Desmond & Glover, 2002).

The first level analysis calculates parameter estimates for the contrasts of interest within each subject and for each run (FEAT, M. W. Woolrich et al., 2001; fixed effects, no slice timing correction, see Perrachione and Ghosh, 2013). The calculation convolves the stimulus time course with the double gamma model of the HRF (Glover, 1999), which is a standard approach in fMRI analysis and recommended for sparse sampling sequences (Perrachione & Ghosh, 2013). The result of the convolution models the expected BOLD response (Figure 11).

The two passes of higher level analysis compute a group level result (FEAT, v6.00, FSL; M. Woolrich et al., 2004). The first higher level analysis (FEAT; fixed effects), combines the two runs within subject. This pass processes data from all participants in a single computation modelling the mean for each participant. The final higher level analysis (FEAT; mixed effects) combines the results from individual subjects into a group average. A voxel-wise threshold for all top level analyses (z=3.1 and α=0.001, see, Woo, Krishnan, and Wager, 2014) filters the activation maps, and among voxels with values that exceed the threshold, cluster-wise probabilities control the family-wise error rate and determine statistical significance of (α=0.05, p=0.0125, see Eklund, Nichols, and Knutsson, 2016; Woo et al., 2014; corrected for multiple comparisons, Judd, McClelland, and Ryan, 2009; Worsley, 2001). To probabilistic atlases, the Harvard-Oxford Cortical Structural Atlas, which was provided by the Center for Morphometric Analysis (CMA) at Massachusetts General Hospital, and the Jülich Histological Atlas, which was developed by Zilles and Amunts and provided by Simon Eickhoff, guide identification of anatomical regions of the brain with which to align clusters that have significant functional activation.

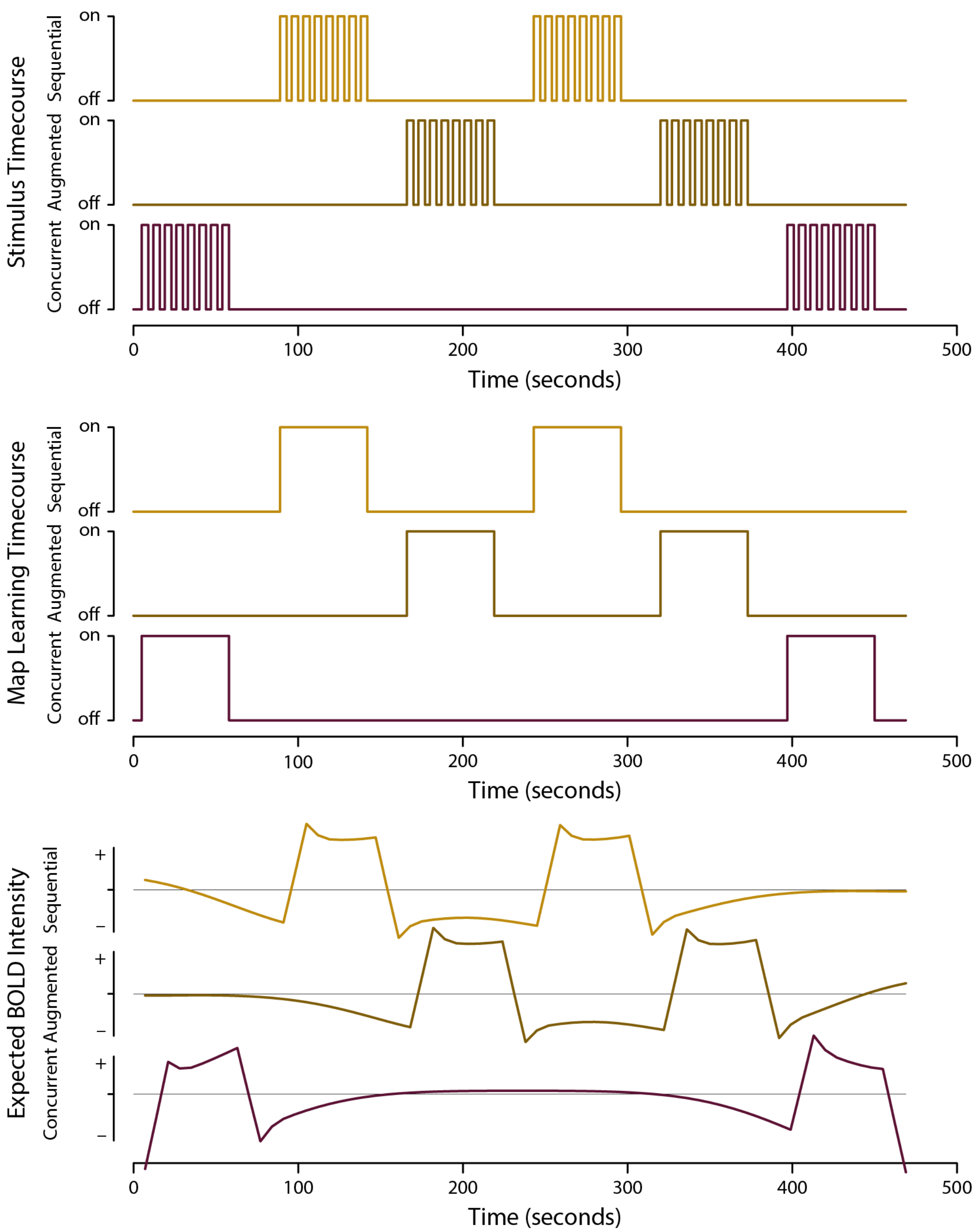

FIGURE 11. The analysis of neuroimaging data measures correspondence between the observed BOLD signal and a model of the expected brain activity. The auditory stimulus occurs in short bursts during the quiet interval between image acquisitions (top). Interpretation of the display contents spans multiple bursts of the auditory stimuli; grouping the the individual events reflects the sustained activity (middle). The model convolves the stimulus time course with a double gamma approximation of the hemodynamic response function (HRF; bottom).

CHAPTER IV

RESULTS

Twenty two participants[1] (ages 20 to 62, average = 34 years old; 16 females) completed paired training and scanning sessions. Demographic data provide information about the sample population (see Table 4). Participants self-reported level of music experience varies widely from no experience (five participants) to one participant who has 41 years of experience. The majority of participants use maps regularly, where map use includes the use of direction-finding applications on mobile devices. Although recruiting does not explicitly select for dominant hand, all participants are right handed. Of the sample population, twenty-one participants gave consent to include data collected during their sessions in a public release.

Training session durations range from 34 to 63 minutes (average = 51 minutes; median = 52 minutes). The session duration includes time spent describing the study and conducting the informed consent process. The planned one-hour schedule provides sufficient time to complete the majority (16 of 22) of the scanning sessions; none of the scanning session durations exceed 75 minutes. As needed, the additional time accommodates, e.g., providing clarification on the MR safety screening form and adjusting a set of MR safe glasses for participants to wear in the scanner.

[1] With fixed funding to support data collection, two situations produce an overall study population size that is smaller than the 24 participants in the original plan. One participant chose to discontinue participation before completing the scanning session. One late cancellation was a billable scanning session, for which there is no data.

TABLE 4. Aggregated demographic details show diverse levels of experience for the sample population who participated in the study sessions.

Age (years)	
Range	20–62
Average	34
Median	28

Gender	
Female	16
Male	6

Preferred Hand	
Right	22

Music Experience (years)	
Range	0–41
Average	9
Median	5

Map Use	
Daily	14
Occasionally	7
Rarely	1

Behavioral Results

Response time and accuracy measure performance during the scanning session. The analysis describes only the portion of the task associated with the response activity (for analysis of listening activity, see neuroimaging results, described below).

Overall response times are faster for both the sequential and concurrent map types in contrast to the augmented-sequential map type (Table 5). The response times and variability differ between response condition (Figure 12); under the active listening condition, response times have particularly high variability, ranging from 0.60 seconds to 56.96 seconds (Figure 13). Under the active listening response condition, the observed difference in response time between augmented-sequential and concurrent map types is unlikely to occur by chance (Friedman $T=4.55$, $p=0.03$; Table 6).

TABLE 5. Average response time (top) and accuracy (bottom) values across the 3 map type × 2 response type conditions.

Average Response Time (seconds)

	Recall	Active Listening	Overall
Sequential	4.0	16.8	10.4
Augmented-sequential	4.1	25.8	15.0
Concurrent	4.2	12.0	8.1
Overall	4.1	18.2	11.2

Average Accuracy (%)

	Recall	Active Listening	Overall
Sequential	57.1	85.7	71.4
Augmented-sequential	81.0	71.4	76.2
Concurrent	35.7	64.3	50.0
Overall	57.9	73.8	65.9

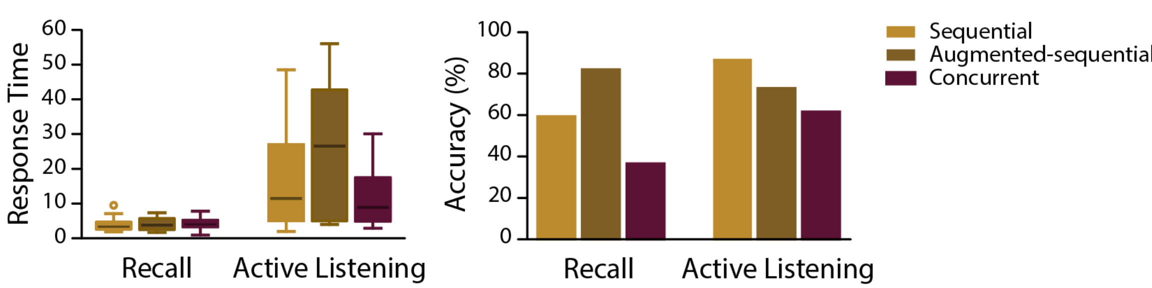

FIGURE 12. Behavioral data, in the form of response time and accuracy, provide descriptive measures of task performance. Within each participant, the response time value represents the average of two trials for each condition.

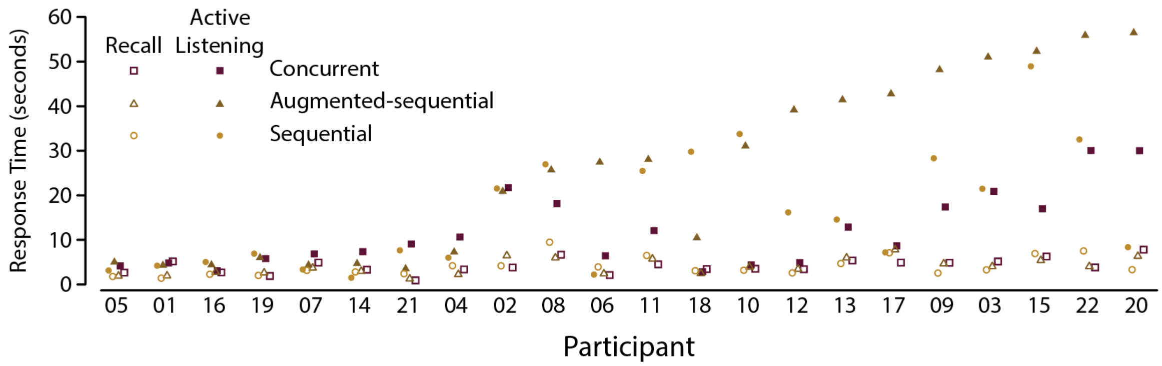

FIGURE 13. Response time by participant, sorted on the maximum response time (for any trial condition). Within each participant, the value represents the average of two trials for each condition. In contrast to an aggregate metric, this response time plot makes visually apparent the high variability of response time (particularly in response to the augmented-sequential map type).

TABLE 6. The Friedman Rank Sum test for differences in means is applied to evaluate the influence of map type on response time.

	Recall		Active Listening	
Contrast	T-score	P-value	T-score	P-value
Sequential vs. Augmented-sequential	1.636	0.201	2.909	0.088
Augmented-sequential vs Concurrent	0	1	4.546	0.033

Within the active listening condition, there was expected to be a relationship between the response time and the onset of the second target region as playback of the second target region provides task critical information. Regardless of response condition, participants respond both before and after the display sonifies task critical data (Figure 14). Responses that occur before the onset of the second target region either rely on a mental representation or indicate guessing.

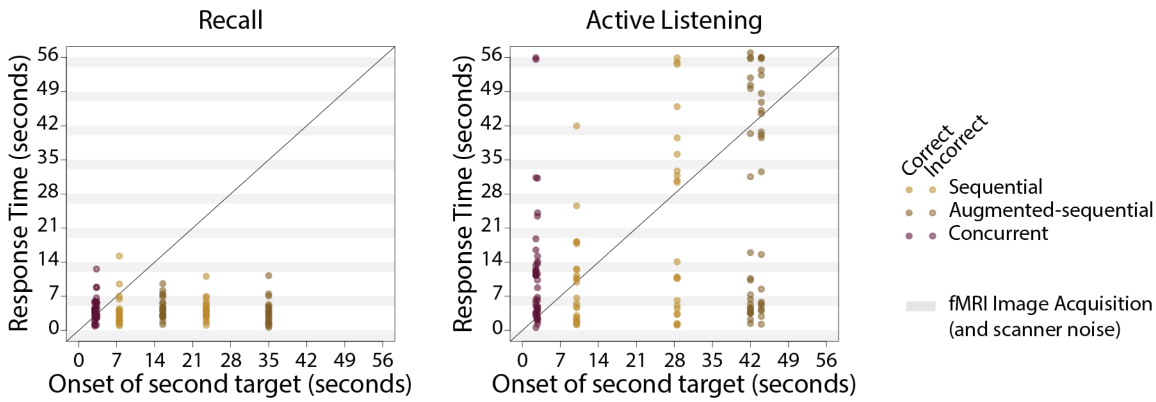

FIGURE 14. The relationship between response time and the onset of the second target region suggests different strategies across the two response conditions and within the active listening condition. Response times corresponding to judgments made from a mental representation of the map contents (*recall* condition) are similar across all map types and tend to occur within the first TR. A bimodal distribution of response times is apparent among observations from the sequential playback (particularly those in which the second target region played more than halfway through the map playback) suggests that participants choose either to make a judgement from memory or to wait until the display sonifies task-critical information.

TABLE 7. A McNemar test for significant changes evaluates the influence of map type on accuracy for the two sets of map stimuli, which were presented in a counterbalanced order.

		Recall		Active Listening	
	Contrast	χ^2	P-value	χ^2	P-value
Set A	Concurrent vs. Augmented-sequential	7.692	0.006	0	1.000
	Augmented-sequential vs. Sequential	0.444	0.505	0.1	0.752
Set B	Concurrent vs. Augmented-sequential	4.267	0.039	2.5	0.114
	Augmented-sequential vs. Sequential	5.143	0.023	2.25	0.134

The concurrent map type has the lowest accuracy scores and the augmented-sequential map type has the highest accuracy scores (Table 5, above). Surprisingly, accuracy in response to the augmented-sequential map type is higher for recall than for active listening (81.8% and 72.1%, respectively).

Neuroimaging Results

The contrast between the concurrent and augmented-sequential map type produces one statistically significant cluster in which activation in response to the concurrent map type was stronger than that to the augmented-sequential map type (Table 8 and Figure 15). The statistically significant cluster overlaps with the left planum temporale and auditory cortex; although only the result on the left side is statistically significant, the data show bilateral differences in activation.

Three statistically significant clusters indicate a stronger response to the augmented-sequential map type than to the sequential map type (Table 8 and Figure 16). Each statistically significant cluster is unilateral, but the data reveal bilateral activation corresponding to each of those clusters that passes the voxel-wise threshold but does not survive the cluster-wise correction for family wise error.

TABLE 8. Two directional contrasts reveal statistically significant differences between the neural activation in response to the concurrent map type and that of the augmented-sequential map type, and between the augmented-sequential and sequential map types. The location of the cluster is described by the coordinates (MNI152) of the peak z-statistic.

Contrast	Cluster	x	y	z	voxels	p-value
Augmented-sequential > Sequential	1	-44	-24	14	655	<0.001
	2	50	-26	12	403	<0.001
	3	-6	-74	2	236	<0.001
	4	16	-56	2	127	0.006
	5	40	2	-12	121	0.008
Concurrent > Augmented-sequential	1	62	-18	8	1005	<0.001
	2	-46	-26	10	773	<0.001

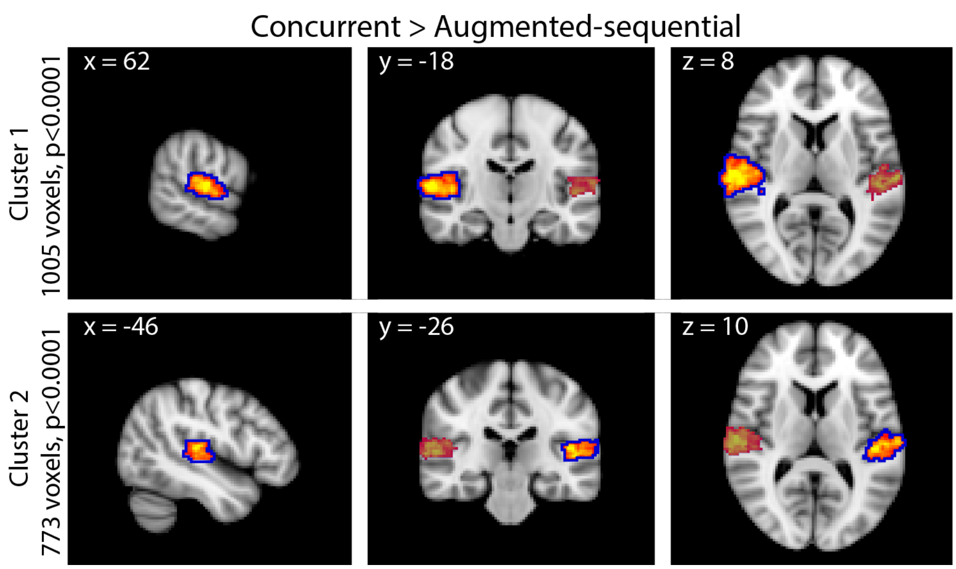

FIGURE 15. A cluster-wise analysis finds statistically significant clusters in the left and right auditory cortex. The contrast in this analysis relates the level of neural activation to the temporal arrangement of the auditory geographic map symbology: activation in response to the concurrent map type is greater than that of the augmented-sequential map type. The significant cluster is outlined in blue, and the image coordinates correspond to peak activation within the cluster.

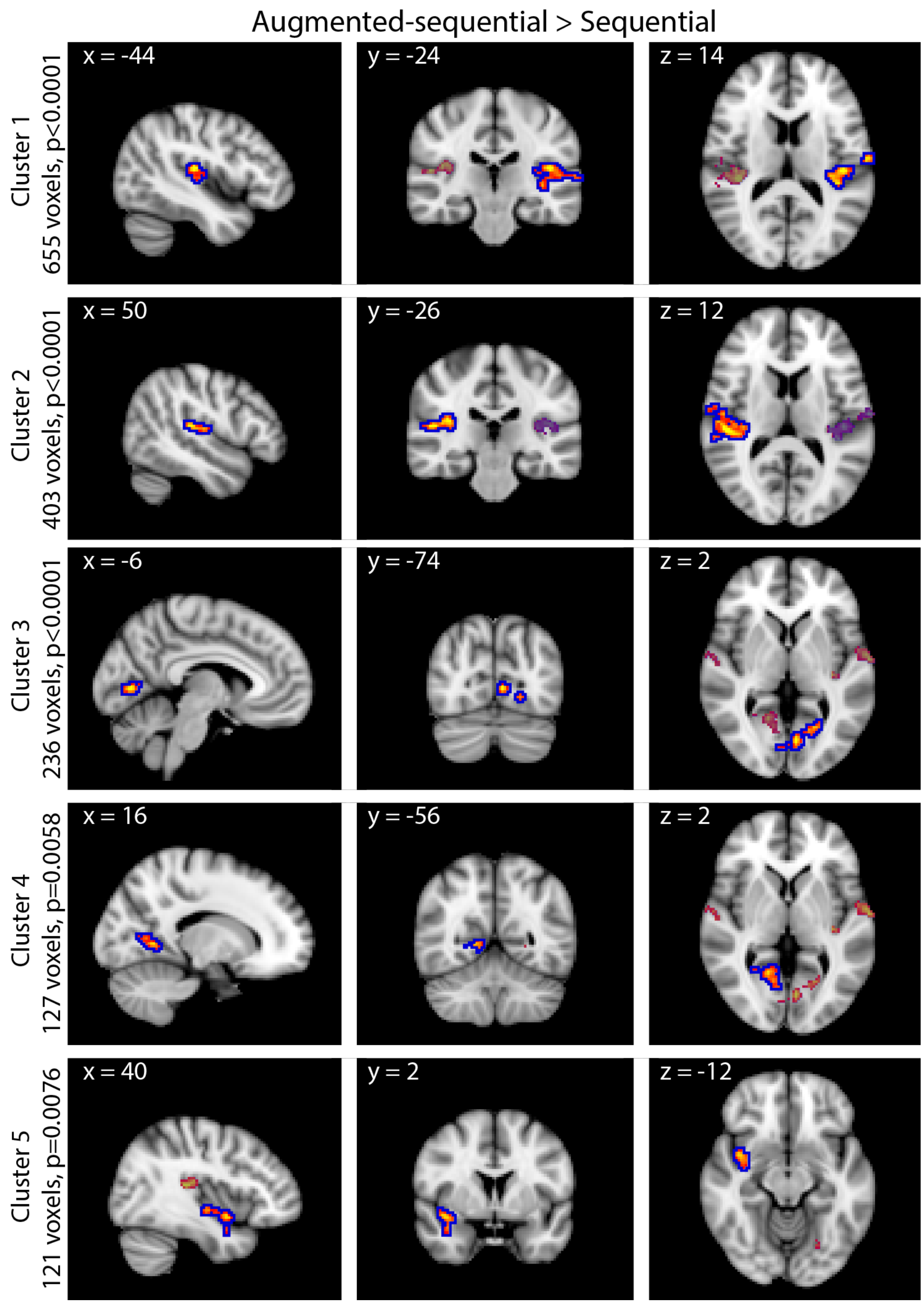

FIGURE 16. A cluster-wise analysis finds five statistically significant clusters that overlap with the left and right auditory cortex (clusters 1 and 2), left and right visual cortex (clusters 3 and 4) and the right insula (cluster 5). In each cluster, neural activation corresponds to the level of encoded information: activation in response to the augmented-sequential map type is greater than that of the sequential map type. The significant cluster is outlined in blue, and the image coordinates correspond to peak activation within the cluster.

Design Feedback

Participant comments during the training session provide insight into user preferences and usability related to the three auditory map types. This summary groups comments according to three themes: mapping data to the dimensions of audio, reflection on the three audio types, and reflection on the map reading task.

Mappings from data to audio parameters receive mixed responses. One participant notes that the "pitch and rhythm easy to follow; they have order" (sub-04). But, in contrast, others find the differences in speed difficult to interpret. The "warble" (sub-12) of the notes in the east distorted the loudness. And, because there are "more notes [in the east,] its hard to tell where one cell ends and the other begins" (sub-03). The use of volume to convey data values also requires practice. It was initially hard to "tell what is loud" (sub-17) and "volume in different peaks is hard to compare" (sub-05). Reflecting on their strategy for listening to the maps, one participant reports that he is "figuring out louder locations, but comparing peaks was harder" (sub-18).

Overall, participants agree that the sequential and augmented audio types are easier to interpret than the concurrent audio. One participant describes the map display that follows a reading order saying: "I like these [sequential] better, they're not as busy" (sub-16). But despite the simplicity, participants note that the sequential audio type is also challenging ("eight series are hard to remember," sub-01, and "all the same tones [make it] way harder to delineate between rows," sub-12) or inefficient ("obviously took longer but, information is just in one [dimension]," sub-09). Another participant remarks that it is "hard to notice all three [frequency, note duration, and volume] at the same time" (sub-04) in the concurrent audio type.

Beyond the parameter mapping and types of audio maps, the data itself also influenced perceived difficulty. One participant reflected on the two instances of the concurrent audio maps in the training session, saying of the first one "it was so obvious," but of the second one "this is tricky" (sub-20). Despite an attempt to control map complexity in the stimulus design, the spatial patterns in the data still influence the ease of interpreting map content.

CHAPTER V

DISCUSSION

This research seeks to better understand the role and influence of temporal arrangement of an auditory geographic map display. The analysis of data from the empirical study provides evidence of differences in behavioral and neural responses to the three auditory map types. The discussion reflects on the auditory-geographic-map design, provides an interpretation of the empirical results, and considers potential threats to validity.

Reflection on Audio Design

The sequential and augmented-sequential map types create a one-to-one mapping between data records and sound events. But, by stringing out the data into a linear sequence, the map playback duration is proportional to the size of the dataset. The linear-time-based presentation was feasible for the small test data set, but the approach scales poorly to larger, real-world raster data sets. Long sequences place a high demand on working memory and require listeners to mentally reconstruct two-dimensional space. Further, compared to visual maps, which convey spatial patterns in a few seconds, communication that takes almost a minute to render (for even a small 8×8 raster data set) is prohibitively inefficient. Revisiting the established convention of displaying space over extended periods of time prompts investigation of novel symbology for auditory geographic maps.

The auditory map symbology revisits the design decisions around the temporal aspect of the auditory display. This research uses functional magnetic resonance imaging (fMRI) to evaluate the cognitive impacts of the new types of auditory maps, and the fMRI acquisition sequence constrains the duration of audio playback. Timing in the map display had to align with short (five second) silent intervals between bursts of scanner noise. The design also prioritizes the isolation of temporal order and map duration over polished aesthetic properties of the maps. Within these constraints of the broader research project, the auditory dimensions of frequency (pitch) and note rate (tempo) create a two dimensional reference frame.

The concurrent auditory symbology achieves the goal of reducing geographic map playback duration, but is sub-optimal. There are certainly more aesthetically pleasing approaches, but this version conforms to requirements of the fMRI scanning environment and its production can be fully automated.[1] The original notion of a concurrent auditory geographic map symbology leveraged musical structures of sound. A combination of multiple frequencies that become a chord is present in the concurrent map type as implemented. However, the second dimension of audio that aligns with the second spatial dimension of the data reflects a compromise in favor of functionality that is both available in the sound synthesis software and feasible within the scope of this project.

When listening to the three auditory map types, the listener has no control over the presentation order or speed. The sequential map type is the most prescriptive: the display presents one data value at a time in a fixed order and with fixed note duration as determined by the cartographer. In contrast, the concurrent map type relinquishes some control over the map-listening experience. By presenting multiple pieces of information at any given time, the concurrent map type allows the listener to attend to various aspects of the complex sound as they

[1] In the long run, a fully automated approach is neither appropriate nor possible for geographic map design. Just as experienced cartographers modify the automated symbology output from GIS software such as ArcMap using graphic design software such as Adobe Illustrator, auditory geographic maps will also require manual polishing.

choose. For comparison, a static visual geographic map is not typically considered "interactive", but the viewer controls the location of their gaze, attending to and gather detailed information from viewer-selected regions of the map in a viewer-specified order. Research in user interfaces explores the ways that interface designers can expose control of the experience to the user (see, e.g., guidelines for interface design by Schneiderman and Plaisant, 2010). A user's ability to change the speed of an animation is an example of how an interface designer or a cartographer may give some control to a user or map reader. However, the software still mediates the interaction; the user can only control those aspects of the display that the interface designer choses to expose. In the case of the concurrent map type, listeners might selectively attend to low frequency notes as the map plays, but the current implementation does not provide an option to change the specific frequencies or the speed of the playback.

Interpretation of Empirical Results

The behavioral data addresses the first research question (RQ1: influence of temporal aspect of audio on communication of spatial patterns). As a tool to communicate spatial pattern, and as measured by response accuracy, this concurrent auditory map type is ineffective. The difference in accuracy in the contrast between concurrent and augmented sequential map types under the recall response condition is statistically interesting (McNemar χ^2=7.69, p=0.006; Table 7). However, the inconsistency between the two sets of map stimuli indicates that the observed difference may not be practically significant. While overall accuracy of responses to trails under the concurrent-map-type condition, post-hoc

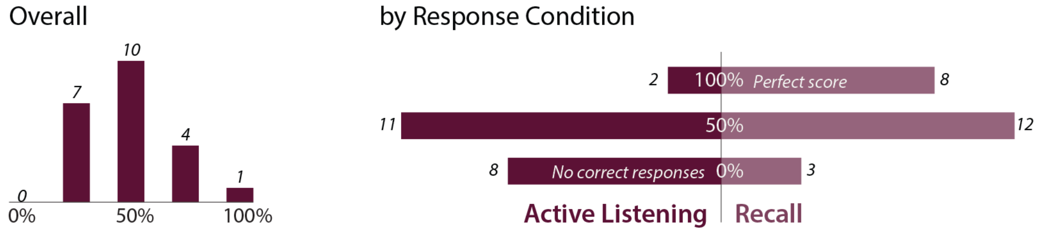

FIGURE 17. Performance accuracy for the concurrent map type is at the level of chance (left). Response condition may modulate performance accuracy (right), but the exploratory visualization and small number of observations precludes drawing a robust conclusion.

exploration of the accuracy data suggests that an interaction may exist between accuracy and the response condition (recall or active listening; Figure 17).

While no differences in response time between the paired contrasts of interest stand out as being statistically interesting (Table 6), the notable differences in response time across the response conditions (recall and active listening; see Figure 12) indicate that at least two map use cases should be considered separately in future analysis. Patterns in the observed response time values also point to potential differences in task strategy.

In the sequential and augmented-sequential map types, response time observations suggest more than one strategy.[2] With the playback of the sound events that represent the second response option, the auditory display provides sufficient task-critical information to make or confirm a judgement about relative value. In the active listening response condition, some participants seem to rely on a mental representation of the data and enter their response shortly after the response prompt appears. Others apparently rely on the playback (either a literal

[2] The number of strategies may also increase with the inclusion of a more diverse participant population. For example, Pasqualotto, Spiller, Jansari, and Proulx (2013) report that people who are blind tend to adopt different behavioral and cognitive strategies for handling spatial data than their sighted peers.

repetition of the auditory map representation or mental rehearsal[3] to extract sufficient information to make a decision or to confirm the result of querying a mental representation. Interestingly, the direction of the difference in accuracy for the sequential and augmented-sequential map types differ between the two response conditions. In contrast to a previous report that demonstrates an increase in accuracy with the encoding of more information in a complex sound (Schito & Fabrikant, 2018), accuracy in response to the augmented-sequential map type is lower than that of the sequential map type under the active listening response condition. While interesting, this directional relationship could be an artifact of the small sample size rather than a true difference in mean accuracy. Across all the map types, performance is likely to improve with additional training and practice that provide familiarity with both the auditory representation and alignment between the auditory geographic map and the visual response prompt.

The neuroimaging data addresses the second research question (RQ2: contrast in neural activity in response to sequential vs. continuous auditory stimuli). Through a cluster-wise analysis, the data reveal statistically significant differences in activation associated with both the level of information explicitly encoded in sound events and the temporal arrangement of the sound events.

The two contrasts between the augmented-sequential and sequential map types provide measures of the influence that the level of encoded information has on listeners' neural response. In the contrast that measures differences in neural activation in which the response to the augmented-sequential map type is greater than that of the sequential map type (augmented-sequential > sequential),

[3] Nees and Walker (2008) describe a similar temporal delay proportional to the length of time between the start of the auditory representation and the display of task critical information as indication of a "sensory-musical encoding," describing the encoding strategy as being "like a tape recorder in their minds"(Nees & Walker, 2011).

five groups of voxels pass the threshold of statistical significance using cluster-based analysis to control the family-wise error rate (no groups pass the voxel-wise threshold for the opposite contrast in which activation in response to the sequential map type is greater than that of the augmented-sequential map type). Bilateral activation in auditory cortex (clusters 1 and 2) is not surprising. The activation on the left side (cluster 1) is somewhat superior to (toward the top of the head from) the auditory cortex, which, according to the Juelich Histological Atlas, falls in the secondary somatosensory cortex.[4] Analysis reveals bilateral activation in visual cortex in clusters 3 and 4. A response in the visual cortex to auditory stimuli is interesting, but again not unexpected. The difference in activation could stem from many causes ranging from a cognitive strategy for handling spatial information that is biased by visual experience, to overwhelming difficulty of the task that recruits help from expansive regions of the brain. Cluster five overlaps with the right insula. Reports of asymmetry in neural responses are common (e.g., asymmetry in the planum temporale among musicians Keenan, Thangaraj, Halpern, & Schlaug, 2001). And allocation of non-verbal auditory processing is one of many reported functions of the insula (e.g., auditory sequencing, see Bamiou, Musiek, & Luxon, 2003).

The two contrasts between the concurrent and augmented-sequential map types measure differences in brain response attributable to temporal arrangement. The contrast that identifies voxels that exhibit a stronger response to the concurrent map type than the (concurrent > augmented-sequential) produces to groups of voxels that pass the threshold for statistical significance using cluster-

[4] The supplementary motor area, located in somatosensory areas has been reported to respond more strongly to temporal than spatial stimuli, Coull, Charras, Donadieu, Droit-Volet, and Vidal, 2015; but the reported location is superior (toward the top of the head) to the peak activation observed in this study.

based analysis to control the family-wise error rate (the contrast in the opposite direction produces no statistically significant clusters). Statistically significant results of the contrast between the concurrent and augmented-sequential map types fall bilaterally in auditory cortex in clusters 1 and 2. With different levels of auditory information to process at a given time, differential activation in the primary sensory areas are not surprising.

Contrary to a-priori expectations, the data do not show a difference in activation within the supplementary motor area (SMA), which contributes to the accumulation of information over time (Coull et al., 2015). Nor are there apparent differences in activation in areas of the brain known to process spatial information (e.g., parahippocampal place area, Epstein, Harris, Stanley, and Kanwisher, 1999, which is active when viewing maps, Rozovskaya and Pechenkova, 2012; or the retrosplenial cortex, Auger et al., 2012). In future work, the use of a region-of-interest analysis could help explain how the auditory stimuli influence these task-relevant brain regions.

Reflection on Experimental Design and Threats to Validity

The auditory-map designs, which satisfy requirements for experimental evaluation in an fMRI environment, incorporate many decisions that influence the interpretation of results, but their full exploration was beyond the scope of this study. One such decision was the use of a small raster data set as the underlying geospatial data. The simple 8×8 raster grid is an abstract form of geospatial data,[5]

[5] Real world geospatial data sets tend to be much larger, and future work will need to investigate the extent to which results obtained from small, low resolution data sets scale and apply to larger and more complex data sets. While data simplification is a common and necessary step in cartographic design (Muehrcke, 1973), there are many possible approaches. Extending the ideas about auditory geographic map design from this research to more realistic applications could involve implementing the three auditory map types for larger raster data sets. Or it could involve

and could easily apply to or be interpreted in a non-geographic context. As a common type of geospatial data, the use of population data bolsters face validity by providing geographic context for the experimental task (similar to the way that Langers et al., 2005 used a behavioral task to monitor subjects, but did not include behavioral data in their analysis).

As with any user interface, there are inconsistencies between subjective impressions of effectiveness and quantitative metrics of performance. Some of the participants approached the auditory symbology with initial skepticism; however, with just a limited amount of practice (less than one hour), participants were able to make some sense of the auditory map displays. Participants occasionally express concern that they were performing poorly, but overall accuracy was better than chance. Although a high level of performance could happen by chance, there is also a possibility that the novelty of *listening* to maps negatively impacted their confidence. As an example, on participant stated "I'm probably tone deaf" (sub-09), yet that participant responded correctly to three of the four trials of the concurrent map type (in which discriminating between frequencies is critical to the task).

Several potential threats to validity limit the evaluation and conclusions that can be drawn from the behavioral data. First, while behavioral testing illuminates trends in performance, there is high variability between individuals. Demographic data provide clues about prior experience with active music listening, they cannot perfectly capture all aspects of an individual's history that could influence performance. Differences between subjects' performance, which are attributable to previous experience with map reading (Muehrcke, 1973, related

a conceptual shift in the way that structural features of the data set are described and handled, e.g., an object-based approach that leverages hierarchical reasoning (Timpf & Frank, 1997).

to visual map reading) or musical training (weak, but statistically significant, correlation with task performance, Schito and Fabrikant, 2018; modulated strength of BOLD response, Coffey, Musacchia, and Zatorre, 2017) likely influence task performance. At the same time, however, the variety of backgrounds and experience help break free from design conventions (see Montello, 2002; Robinson, 1986) that may be introducing or propagating bias in development of sonification techniques in cartography. While the diversity of participants and their prior experience introduces noise into the measurements, detectable differences that remain may indicate universal properties of perception and cognition. Second, the MR scanning environment is a threat to ecological validity. Participants endeavour to listen carefully to subtle changes in audio streams while they lie on their back, with their head in a small (70 centimeter diameter) tube, and the scanner bombards them with loud intermittent noise. Participants mention that the scanner noise distracts from the map listening task at inopportune points in time. Right after hearing an audio burst, information is in sensory memory (Levitin, 2002), and when interleaved volume acquisition starts, the acoustic scanner noise clobbers the task relevant information, disrupting encoding and mental rehearsal. The scanner noise comes in just as participants are trying to reflect on the data display and synthesize new information with their mental representation of past information. The extent of the influence of the MR scanning environment on performance of the map listening task is unknown. Future work that conducts more trials in a computer-lab setting could provide a reference data set to evaluate the impact that the the scanning environment has on task performance. Finally, the duration of each map limits the number of trials within the one-hour sessions. The small number of observations yield low statistical power. Still, descriptive summary statistics can provide insight into differences between the auditory map types and inform interpretation of the neuroimaging data.

While neuroimaging offers a unique perspective in the assessment of map effectiveness, challenges and threats to validity accompany the method. First, the extreme simplification of the map data and the unusual environment for map listening threaten ecological validity. The simple map data occupy a square grid made up of eight rows and eight columns (8×8 raster grid, however, “real maps that are almost never composed of convenient little squares” Olson, 1975). But raster data facilitate experimental control, and small data size constrains playback duration for sequential map symbology. Providing even twice the isotropic resolution in the data would quadruple playback duration, and would tax the listener’s working memory even further. Second, participants perform the map listening task in the MR scanner, which allows acquisition of brain images. The MR scanning environment likely influences behavioral performance through distraction (Liem, Lutz, Luechinger, Jäncke, & Meyer, 2012) and added stress associated with the unfamiliar space. And acoustic noise of the scanner during acquisition leads to activity in the auditory cortex that is unrelated to the task. Sparse sampling mitigates saturation of the auditory response and makes auditory stimuli easier to hear. An unfortunate side effect of the sparse sampling is intermittent loud noise. The scanner noise creates an abrupt interruption just as participants are trying to process what they have just heard. Third, even though the design of the auditory symbology endeavors to isolate the temporal component of the auditory map implementations, the degree to which a single aspect of the auditory display can be isolated without sacrificing ecological validity is limited. Recognizing the perceptual interactions between dimensions of audio it was not realistic to expect that a change to the temporal arrangement within an audio stream would not influence other aspects of the audio. A hybrid augmented-sequential auditory map symbology serves as an intermediate step that reduces, but does not eliminate, confounding interactions in the transition between the established sequential playback auditory geographic map type and the more compact concurrent audio representation.

CHAPTER VI

CONCLUSIONS

The three auditory geographic map types introduced in this manuscript provide further evidence that sound can encode multidimensional data, and take a small step toward understanding how to do so in a way that listeners can understand. While performance on trials that use the concurrent map type was not strong, the task established context and directed participants' attention toward interpretation of the auditory stimuli as geographic maps. The neuroimaging, then, provided an alternative way to investigate the influence of temporal arrangement across the three auditory map types.

In response to the research questions posed at the outset of this research, this research identifies patterns in the way that the brain responds to the auditory geographic maps amid imperfections in the specific implementation of the auditory symbology. As implemented in the concurrent map type, the overlapping temporal arrangement of the auditory map symbolization correlates with difficulty interpreting general spatial patterns in the data. Whether attributable to the temporal overlap of symbols or the pairing of frequency and note rate, accuracy on the comparison task is poor for maps using the novel concurrent audio symbology. Even though the specific instance of auditory symbology leaves something to be desired in the realm of cartographic communication, controlled differences between the three map types still provide insight into how design choices influence map listening. Neural activation is stronger bilaterally in the auditory cortex, overlapping the right planum temorale and left Heschl's Gyrus, in response to concurrent audio symbolization of geospatial data in contrast against that of the augmented-sequential symbology. Although not known to selectively process spatial information, stronger activation in the planum temporale suggests future work to better understand the auditory geographic map display as a symbolic language, akin to music or natural language. Further, in contrast to a sequential, single-attribute encoding of data, augmenting the display with additional parameter mappings evokes a greater response bilaterally in the visual cortex. While this activation could indicate recruitment of additional brain regions to handle the more complex sound, implicating the visual cortex in response to auditory stimuli suggests a benefit to using a rich parameter mapping sonification to represent geospatial data.

This dissertation attributes the limited adoption of auditory display in cartography to a lack of established auditory mapping techniques. The ubiquity of visual maps, which provide the majority of experience with maps, feed a visual bias among cartographers that influences not only the types of artifacts that are considered "maps" but also the way that people who are sighted think about spatial information. While these observations may help to explain the existing dearth of auditory maps, they do not justify the continued exclusion of audio from development of new cartographic techniques. To move forward, geographic map design needs to consider audio both early in the process and as a primary display modality. Although many cartographers unfortunately dismiss audio as a potential medium for displaying geographic information, an expansive sonification design space is available for use in auditory geographic maps. Exploring the design space with an ear toward geospatial data facilitates adoption of audio into cartographic design and contributes to the development of diverse geographic map artifacts.

APPENDIX A

DEMOGRAPHIC QUESTIONS

Study participants provide demographic information by responding to a demographics questionnaire. The demographic data describe the sample population participating in the study, but do not contribute to the quantitative analysis. Section IV: Results reports participant responses in aggregate.

1. What is your age?
2. What is your gender?
3. Which hand is your dominant hand?
4. Do you have experience with map reading? If so, how often do you use maps? (daily, seasonally, occasionally, rarely)
5. Do you have music experience? If so, how many years of informal experience or formal training do you have?

APPENDIX B

SCAN SEQUENCE DEVELOPMENT: MULTI-BAND ACCELERATION

The magnetic resonance imaging (MRI) scanner produces substantial noise during image acquisition, and this noise poses a substantial challenge to the presentation of auditory stimuli. Recent techniques that use custom scan sequence parameter promises theoretical advantages to accommodate presentation of the auditory geographic maps. Sparse sampling and multi-band acceleration are two such techniques that balance overall scan duration, acquisition time for a single volume, and spatial resolution of the resulting images while accommodating auditory stimuli. Sparse sampling inserts a quiet interval between acquisitions, during which auditory stimuli can be presented without interference from the scanner noise. Multi-band acceleration reduces the amount of time required to acquire a single volume. This appendix describes the scan sequences and illustrates an observed negative consequence of combining the two techniques.

Multi-band acceleration uses concurrent excitation and readout of two or more slices within a single volume (Larkman et al., 2001). Recording multiple slices at the same time reduces the total time required to collect a volume (TA). The resulting time savings translates into an increase in the number of volumes in a scanning session, a reduction of the scanning-session duration, or an increase in spatial resolution. Images recoded using multi-band acceleration are susceptible to spurious detection from signal bleed across concurrent slices; the impact is minimal for an acceleration factor of two and increases for higher acceleration factors (Todd et al., 2016, evaluated within visual and motor cortices). Combining sparse-sampling with multi-band acquisition offers the benefit of a longer silent

TABLE 9. Custom scan sequence parameters accommodate scanner noise on the presentation of auditory stimuli using both sparse sampling and multi-band acquisition (A). Adjustments to mitigate the observed stripe artifact include slice acquisition order (B), slice spacing (C), and omission of the multi-band acquisition at the expense of spatial resolution in slice thickness (D).

	TR (ms)	TE (ms)	Slice Order	Thickness (mm)	Slice Spacing (mm)	Slice Duration (ms)	Number of Slices	Voxel Size (mm)
A	7000	25	interleaved	2	2	26.7254	72	2×2
B	7000	25	descending	2	2	26.7254	72	2×2
C	7000	25	descending	2	3	39.4681	48	2×2
D	7000	25	descending	4	4	63.8333	31	2×2

interval while maintaining a fixed TR and without sacrificing spatial resolution (De Martino et al., 2015, evaluated within Hershel's Gyrus).

Although in theory the combination of sparse-sampling with multi-band acquisition (Table 9) gives the best of both worlds, pilot data reveals a spin history effect that is visually apparent in the coronal and sagittal views (Figure 18). Data collected under the theoretically grounded parameters reveals an unanticipated artifact, which, in hindsight, is not surprising. The implementation of sparse sampling in this study turns off both the gradient and RF pulses during the quiet interval. The absence of the RF pulses allowed extended relaxation time between TRs. As a result, the first samples (slices) of an acquisition recorded higher intensity values than the later samples in the same acquisition. The multi-band acceleration distributes the first samples across the z-axis of the volume, placing them spatially adjacent to samples that are collected after the application of RF pulses resumes. Ultimately the detrimental impact of the artifact outweighs the apparent advantages of the combined sparse-sampling and multi-band sequence.

The scan sequence parameters used in the study employ sparse sampling, but exclude multi-band acceleration. Forgoing the multi-band acceleration mitigated the stripe artifact, or abrupt changes in signal intensity across slices, but obviously does not eliminate the underlying cause: spin history.[1] Without the acceleration of multi-band imaging, a concomitant aspect of the sequence had to reactively change. As the desired time course of the stimuli restricts adjustments to the acquisition time (TA) or repetition time (TR), the selected scan sequence parameters sacrifice spatial resolution (4 mm slice thickness, in contrast to the originally planned 2 mm).

The option to use headphones is a redeeming side-effect of the decision to forgo multi-band acceleration. Multi-band acceleration necessitates use of the 32-channel head coil, which has relatively restrictive internal dimensions and requires use of insert earphones[2] to present auditory stimuli. Without the need for multi-band capabilities and accepting lower spatial resolution (4 mm slices) in the digital images obtained, the 20-channel head coil, which has more generous internal dimensions, is an option. This larger head coil provides room for MR-safe headphones[3] that offer better sound quality than the insert earphones.

[1] Mitigation techniques such as clustered acquisition with silent steady state (Schwarzbauer, Davis, Rodd, & Johnsrude, 2006), or read-out omissions (Bartsch et al., 2007) could address the spin history effect. Although scan sequence development considered these approaches, clustered acquisition was not available, and an option to apply continuous RF pulses during the quiet interval was not identified within the LCNI supported software until after data collection had concluded.

[2] Sensimetrics, Model S14, http://www.sens.com/products/model-s14

[3] NordicNeuroLab, AudioSystem, http://www.nordicneurolab.com/products/AudioSystem.html

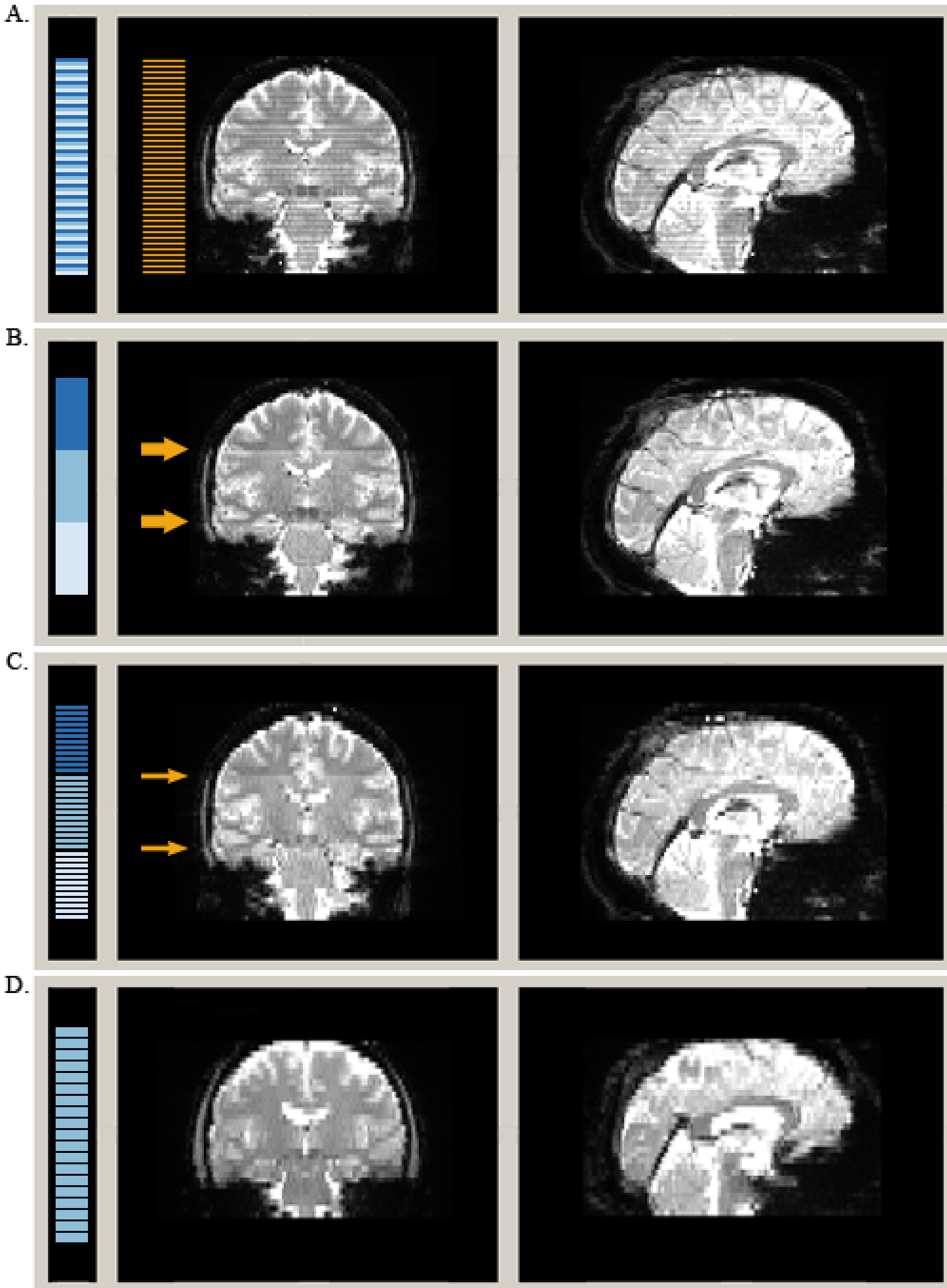

FIGURE 18. An intensity artifact of the multi-band acquisition was visually apparent as a stripe in the coronal (left) and sagittal (right) views. A. Interleaved multi-band acquisition, 2 mm slices; B. Sequential (descending) multi-band acquisition; C. Sequential (descending) multi-band acquisition with a 100 mm gap between slices; D. Sequential (descending) acquisition without multi-band, 4 mm slices.

APPENDIX C

DATA DICTIONARIES

The test instrument produces two types of tabular data:[1] demographic data and behavioral responses. This appendix provides data dictionaries that describe the structure and contents of the tabular data (see also the associated metadata files in JSON format that is available with the data). The data format follows recommended organization for tidy data (Wickham, 2014) and the specific fields they contain conform to the BIDS standard (Brain Imaging Data Structure, v1.0.2). The data are available from OpenNeuro (https://openneuro.org/datasets/ds001415).

Data files from sessions with a single participant are linked by a unique participant identifier: `pNum`. Each participant receives a temporary numeric identifier when the training session is scheduled; the number is a number selected from a uniform distribution (R, `stats::runif`) to obscure the date of participation and protect participant identity. After data collection, all data sets are sorted by the temporary numeric identifier and assigned an archival participant identifier using a one-up counter. The `pNum` is a two digit number with the prefix "sub-" (see the BIDS specifications for requirements related to the subject identifier).

Demographic Data

Participants respond to the demographic questions (see Appendix A: Demographic Questions) on paper at the start of the computer-based training session. The paper

[1] Metadata associated with the neuroimaging data are available in JSON formatted files that conform to the BIDS standard (Brain Imaging Data Structure, v1.0.2).

format allows participants to enter any value they chose for each question, with one exception: participants were asked to select one of the listed levels of experience with map reading. The responses are manually digitized for aggregation and storage in computer files (Table 10); in the case of categorical variables, the actual responses written on paper are listed in the right-hand column. Each participant session contributes one row to a tidy (Wickham, 2014) tabular file of demographics data.

Behavioral Data

The PsychoPy experiment records behavioral responses to the map listening task. Each behavioral data file is associated with the session in which it was collected by the `pNum` in the filename (see the BIDS specifications for details of the file naming convention). A custom script extracts pertinent details from the log files and creates tabular data files (Table 11; see the BIDS specifications for requirements related to the record of event details). Each participant generates two behavioral data sets: one collected during the computer-based training session and one for the scanning session. During the training participants learn about the three map types and the task protocol. The resulting data are susceptible to noise. Examples of noise sources include asking questions during the response activity (increasing the response time) or advancing to the response activity before hearing an entire map (skipping the map presentation and then having no basis to respond in the memory condition).

TABLE 10. The data dictionary for demographic data describes the data fields and range of valid values for the digitized demographic data, which were collected through the paper questionnaire.

variable	type (units)	digitized values	description	response values
participant_id	character	"sub-<*NN*>"	Unique identifier for the participant concatenating the prefix "sub-" with a two digit number	<*randomized, assigned*>
gender	categorical	"female" "male"	Participant gender, articulated by participant	"FEMALE", "F", "female" "Male", "male"
age	integer (years)	18–65	Participant age at the time of participation	<*number*>; "<*number*> yrs"; "<*number*> years old"
hand	categorical	"right" "left"	Dominant hand, self-reported by participant	"RIGHT", "Right", "right", "R"
has_map	boolean	True False	Flag indicating whether or not the participant indicated map reading experience through self-report	
level_map	categorical	"daily" "seasonally" "occasionally" "rarely" NA	Level of self-reported map reading experience (or NA if not applicable, i.e., has_map is false)	

Continued on next page

Table 10 – *continued*

variable	type (units)	digitized values	description	response values
has_music	boolean	True False	Flag indicating whether or not the participant indicated informal music experience or formal music training through self-report (training, practice, performance) with music	"yes", "15 years" "no", "No"
years_music	integer (years)	1–65 NA	Number of years of self-reported music experience (or NA if not applicable, i.e., has_music is false)	
test_order	categorical	"AB" "BA"	Counterbalanced order in which the stimulus sets were presented	

TABLE 11. The data dictionary for behavioral data describes the data fields and range of valid values for the behavioral data, which were extracted from the log files.

variable	type (units)	values	description
onset	float (seconds)	2.5–695.5	Stimulus onset relative to the first trigger pulse
duration	float (seconds)	0.0–56.0	Length of time that the stimulus was presented. When the duration is shorter than the duration of the audio clip, the audio played from the start and was truncated when the 'duration' time had elapsed.
trial_type	categorical	"sequential" "augmented" "concurrent"	The experimental condition describing the auditory symbology used in the stimulus
task	categorical	"listen" "response"	The activity participants were asked to complete during each phase of the trial
response_type	categorical	"active" "memory" NA	The activity participants were asked to complete during each phase of the trial (or NA if not applicable, i.e., task is "listen")
map_num	integer	0–18 NA	Map identifier indicating the auditory stimulus
stim_file	character	<*filename*>	Name of the file containing the auditory stimulus waveform (or NA if not applicable, i.e., response_type is "memory")

Continued on next page

Table 11 – *continued*

variable	type (units)	values	description
correct_answer	categorical	"black" "red" NA	The response option which contained the higher data value (correct task response) (or NA if not applicable, i.e., task is "listen")
answer	categorical	"black" "red" NA	The response option selected by the participant in response to the task question (or NA if not applicable, i.e., task is "listen")
is_correct	boolean	True False NA	The response option selected by the participant in response to the task question (or NA if not applicable, i.e., task is "listen")
sparse	float (seconds)	5.0–698.0	Stimulus onset shifted to align with the midpoint of the TR (for use in FSL to account for sparse sampling with a TR longer than TA)

APPENDIX D

EXPERIMENTAL STIMULI

A collection of software tools (Table 12) automate data generation of eighteen raster data sets (Figure 19). Table 13 indicates the study session (training or scanning) to which each map belongs and reflects the degree of balance in the response options (Table 13).

TABLE 12. Software tools from R and Python libraries generate map data and render auditory geographic maps. Post-processing uses Audacity to identify and remove any clipping.

Name	**Version**	**Description**
`R`	v3.1.1	Synthesize data sets
`classInt`	v0.1-23	assign data values into discrete levels (low, medium, high)
`docopt`	v0.4.5	accept and handle command line arguments
`rgeos`	v0.3-121	convert point patterns into grid representation
`sp`	v1.2-3	define structures to represent and plot spatial data
`spatstat`	v1.47	generate clustered point patterns
`RColorBrewer`	v1.1-2	create color palette for visual representations
`Python`	v2.7	Render audio maps
`Pyo`	v0.7.9	render audio waveforms from matrix and tabular data
`argparge`	v1.1	accept and handle command line arguments
`csv`	v1.0	read tabular playback actions for rendering concurrent map type
`math`	v2.7	compute exponents (Python standard library)
`numpy`	v1.12.0	interpret string input as array of numeric values
`Audacity`	v2.0.6	View and edit waveforms

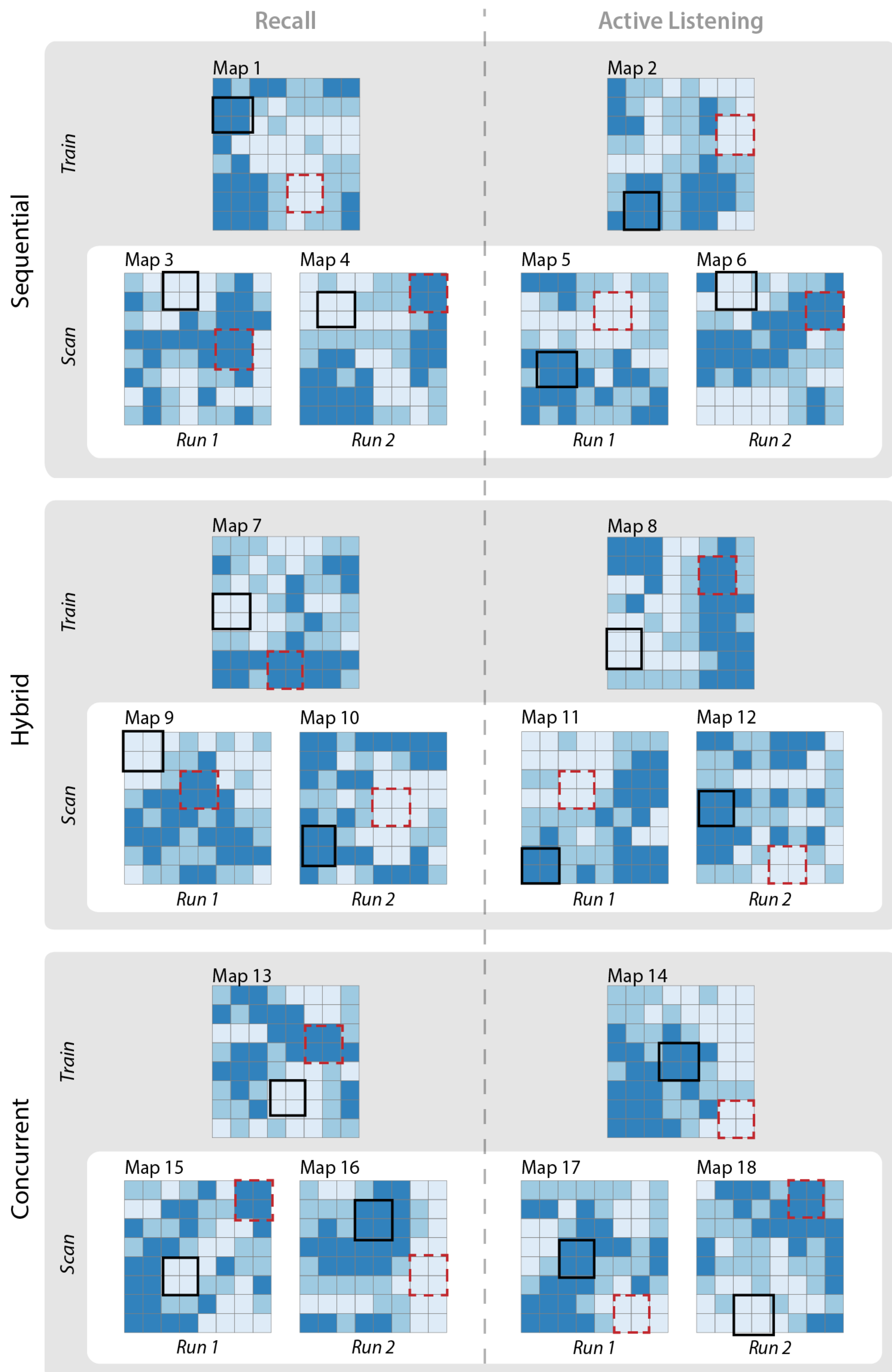

FIGURE 19. A visual depiction of the geospatial data illustrates the map stimuli used in the empirical evaluation. (The visual versions exist for the purpose of illustration in this document; participants do not view the graphics.)

TABLE 13. To the extent possible, the three sets of stimuli balance characteristics of the response option pairs.

	Session		**Assigned Condition**		**Correct**	**Black**		**Red**			
	Type	**Run**	**Audio Type**	**Response**	**Answer**	**X**	**Y**	**X**	**Y**	**Direction**	**Distance**
Map 1	train		sequential	memory	black	-3	2	1	-2	315	5.66
Map 2	train		sequential	active	black	-2	-3	3	1	39	6.40
Map 3	scan	1	sequential	memory	red	-1	3	2	0	315	4.24
Map 4	scan	2	sequential	memory	red	-2	2	3	3	11	5.10
Map 5	scan	1	sequential	active	black	-2	-1	1	2	45	4.24
Map 6	scan	2	sequential	active	red	-2	3	3	2	349	5.10
Map 7	train		hybrid	memory	red	-3	0	0	-3	315	4.24
Map 8	train		hybrid	active	red	-3	-2	2	2	37	6.40
Map 9	scan	1	hybrid	memory	red	-3	3	0	1	326	5.00
Map 10	scan	2	hybrid	memory	black	-3	-2	1	0	27	4.47
Map 11	scan	1	hybrid	active	black	-3	-3	-1	1	63	4.47
Map 12	scan	2	hybrid	active	black	-3	0	1	-3	323	5.00
Map 13	train		concurrent	memory	red	0	-2	2	1	57	3.61
Map 14	train		concurrent	active	black	0	0	3	-3	315	4.24
Map 15	scan	1	concurrent	memory	red	-1	-1	3	3	45	5.66
Map 16	scan	2	concurrent	memory	black	0	2	3	-1	315	4.24
Map 17	scan	1	concurrent	active	black	-1	0	2	-3	315	4.24
Map 18	scan	2	concurrent	active	red	-1	-3	2	3	63	6.71

REFERENCES CITED

Aguirre, G. K. & D'Esposito, M. (2000). Experimental design for brain fMRI. In C. T. Moonen & P. A. Bandettini (Eds.), *Functional MRI* (Chap. 30, pp. 369–380). Springer.

Alexander, R. L., O'Modhrain, S., Gilbert, J., Zurbuchen, T., & Simoni, M. (2012). Recognition of audified data in untrained listeners. *Proceedings of the 18th International Conference on Auditory Display (ICAD2012)*, 135–140. Retrieved from http://hdl.handle.net/1853/44398

Alexander, R. L., Zurbuchen, T. H., Gilbert, J., Lepri, S., & Raines, J. (2010, June). Sonification of ACE Level 2 Solar wind data. In E. Brazil (Ed.), *Proceedings of the 16th International Conference on Auditory Display*. Washington, D.C. Retrieved from http://hdl.handle.net/1853/50049

Alty, J. L. & Rigas, D. (2005). Exploring the use of structured musical stimuli to communicate simple diagrams: the role of context. *International Journal of Human-Computer Studies*, *62*(1), 21–40. doi:10.1016/j.ijhcs.2004.08.003

Amaro, Jr., E. & Barker, G. J. (2006). Study design in fMRI: basic principles. *Brain and Cognition*, *60*(3), 220–232. doi:10.1016/j.bandc.2005.11.009

Asquith, L. (2013, December). Listening to data from the large hadron collider. *TEDx Zurich*. Retrieved from https://www.youtube.com/watch?v=iQiPytKHEwY

Auger, S. D., Mullally, S. L., & Maguire, E. A. (2012, August). Retrosplenial cortex codes for permanent landmarks. *PLoS ONE*, *7*(8), e43620. doi:10.1371/journal.pone.0043620

Baerwald, T. J. (2010). Prospects for geography as an interdisciplinary discipline. *Annals of the Association of American Geographers*, *100*(3), 493–501. doi:10.1080/00045608.2010.485443

Bamiou, D.-E., Musiek, F. E., & Luxon, L. M. (2003). The insula (Island of Reil) and its role in auditory processing: literature review. *Brain Research Reviews*, *42*(2), 143–154. doi:10.1016/S0165-0173(03)00172-3

Bandettini, P. A. (2012, August). Twenty years of functional MRI: the science and the stories. *NeuroImage*, *62*(2), 575–588. doi:10.1016/j.neuroimage.2012.04.026

Bartsch, A. J., Homola, G., Thesen, S., Sahmer, P., Keim, R., C. F. Beckmann, ... Bendszus, M. (2007). Scanning for the scanner: FMRI of audition by read-out omissions from echo-planar imaging. *NeuroImage*, *35*(1), 234–243. doi:10.1016/j.neuroimage.2006.11.026

Baumann, S., Griffiths, T. D., Sun, L., Petkov, C. I., Thiele, A., & Rees, A. (2011, April). Orthogonal representation of sound dimensions in the primate midbrain. *Nature Neuroscience*, *14*(4), 423–425. doi:10.1038/nn.2771

Beckmann, C., Jenkinson, M., & Smith, S. (2003). General multi-level linear modelling for group analysis in FMRI. *NeuroImage*, *20*, 1052–1063. Retrieved from http://www.fmrib.ox.ac.uk/analysis/techrep/#TR01CB1

Belin, P., Zatorre, R. J., Hoge, R., Evans, A. C., & Pike, B. (1999). Event-related fmri of the auditory cortex. *NeuroImage*, *10*(4), 417–429. doi:10.1006/nimg.1999.0480

Berry, B. J. (1964). Approaches to regional analysis: a synthesis. *Annals of the Association of American Geographers*, *54*(1), 2–11. doi:10.1111/j.1467-8306.1964.tb00469.x

Bertin, J. (2011). Analysis of the information. In *Semiology of graphics: diagrams, networks, maps* (W. J. Berg, Trans.) (Chap. 1). Redlands, California: ESRI Press. (Original work published 1967)

Bischoff-Grethe, A., Ozyurt, I. B., Busa, E., Quinn, B. T., Fennema-Notestine, C., Clark, C. P., ... Fischl, B. (2007). A technique for the deidentification of structural brain MR images. *Human Brain Mapping*, *28*(9), 892–903. doi:10.1002/hbm.20312

Boccia, M., Nemmi, F., & Guariglia, C. (2014). Neuropsychology of environmental navigation in humans: review and meta-analysis of fMRI studies in healthy participants. *Neuropsychology Review*, *24*, 236–251. doi:10.1007/s11065-014-9247-8

Bowtell, R. & Peters, A. (1999). Analytic approach to the design of transverse gradient coils with co-axial return paths. *Magnetic Resonance in Medicine*, *41*(3), 600–608. doi:10.1002/(SICI)1522-2594(199903)41:3⟨600::AID-MRM24⟩3.0.CO;2-E

Boyle, K. J., Paterson, R., Carson, R., Leggett, C., Kanninen, B., Molenar, J., & Neumann, J. (2016). Valuing shifts in the distribution of visibility in national parks and wilderness areas in the United States. *Journal of Environmental Management*, *173*, 10–22. doi:10.1016/j.jenvman.2016.01.042

Boynton, G. M., Engel, S. A., Glover, G. H., & Heeger, D. J. (1996, July). Linear systems analysis of functional magnetic resonance imaging in human V1. *The Journal of Neuroscience*, *16*(13), 4207–4221. Retrieved from http://www.jneurosci.org/content/16/13/4207

Brauen, G. (2014). Interactive audiovisual design for cartography: survey, prospects, and example. In D. R. Taylor (Ed.), *Developments in the theory and practice of cybercartography: applications and indigenous mapping* (2nd edition, Chap. 10, Vol. 5, pp. 141–159). Elsevier. doi:10.1016/B978-0-444-62713-1.00010-6

Bregman, A. S. (1990). *Auditory scene analysis: the perceptual organization of sound*. Cambridge, Massachusetts: Bradford Books, MIT Press.

Brewer, C. A. (2016). Labeling maps. In *Designing better maps: a guide for GIS users* (2nd edition, Chap. 6, pp. 103–128). Redlands, California, USA: ESRI Press.

Brewster, S. (2003). Nonspeech auditory output. In J. A. Jacko & A. Sears (Eds.), *The human-computer interaction handbook: fundamentals, evolving technologies, and emerging applications* (Chap. 11, pp. 220–239). Mahwah, New Jersey: Lawrence Earlbaum Associates.

Brewster, S., Wright, P. C., & Edwards, A. D. (1995). Experimentally derived guidelines for the creation of earcons. In *Proceedings of BCS-HCI'95* (pp. 155–159).

Brittell, M. (2018). Seeking a reference frame for cartographic sonification. *International Conference on Auditory Display*. Retrieved from http://hdl.handle.net/1853/60082

Brittell, M., Young, M., & Lobben, A. (2013). The mGIS: a minimal geographic information system accessible to users who are blind. In *Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems* (pp. 554–557). Orlando, Florida: ACM. doi:10.1145/2525314.2525329

Brown, C. & McKinney, J. (2015, February). Accessibility and Section 508. In *Proceedings of the federal GIS conference*. Washington, D.C. Retrieved from http://proceedings.esri.com/library/userconf/fed15/papers/fed_123.pdf.

Brown, L., Brewster, S., Ramloll, R., Burton, M., & Riedel, B. (2003, July). Design guidelines for audio presentation of graphs and tables. In E. Brazil & B. Shinn-Cunningham (Eds.), *Proceedings of the 9th International Conference on Auditory Display*. Boston, Massachusetts. Retrieved from http://hdl.handle.net/1853/50454

Campbell, M. & Greated, C. (1987). *The musician's guide to acoustics*. New York: Schirmer Books.

Caquard, S. & Fiset, J.-P. (2014). How can we map stories? a cybercartographic application for narrative cartography. *Journal of Maps*, *10*(1), 18–25. doi:10.1080/17445647.2013.847387

Cartwright, W. (1996). Hardware, software and staffing requirements of multimedia cartography: an update. *Cartography*, *25*(1), 67–80. doi:10.1080/00690805.1996.9714019

Cassettari, S. & Parsons, E. (1993, March). Sound as a data type in a spatial information system. In *Proceedings of the 4th European Conference on Geographical Information Systems* (Vol. 1, pp. 194–202). Genoa, Italy.

Cherston, J., Hill, E., Goldfarb, S., & Paradiso, J. A. (2016). Sonification platform for interaction with real-time particle collision data from the atlas detector. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (pp. 1647–1653). CHI EA '16. Santa Clara, California: ACM. doi:10.1145/2851581.2892295

Chrastil, E. R. & Warren, W. H. (2014, November). From cognitive maps to cognitive graphs. *PLoS ONE*, *9*(11), e112544. doi:10.1371/journal.pone.0112544

Clingan, I. C. (2017). *Lighthouse*. Encyclopædia Britannica. Retrieved from https://www.britannica.com/technology/lighthouse

Cobb, D. A. (1977, April). Maps and scholars. *Library Trends: Trends in the Use of Library Resources*, *25*(4). Retrieved from https://core.ac.uk/download/pdf/4816264.pdf

Coffey, E. B., Musacchia, G., & Zatorre, R. J. (2017). Cortical correlates of the auditory frequency-following and onset responses: EEG and fMRI evidence. *Journal of Neuroscience*, *37*(4), 830–838. doi:10.1523/JNEUROSCI.1265-16.2016

Cohen, M. S. & Schmitt, F. (2012, August). Echo planar imaging before and after fMRI: a personal history. *NeuroImage*, *62*(2), 652–659. doi:10 . 1016 / j . neuroimage.2012.01.038

Conover, W. J. (1971). *Practical nonparametric statistics*. New York, New York: John Wiley & Sons, Inc.

Constantinidis, C. & Wang, X.-J. (2004). A neural circuit basis for spatial working memory. *The Neuroscientist*, *10*(6), 553–565. doi:10.1177/1073858404268742

Cook, P. R. (Ed.). (1999). *Music, cognition, and computerized sound: an introduction to psychoacoustics*. Cambridge, Massachusetts: MIT Press.

Couclelis, H. (1992). People manipulate objects (but cultivate fields): beyond the raster-vector debate in GIS. In A. U. Frank, I. Campari, & U. Formentini (Eds.), *Theories and methods of spatio-temporal reasoning in geographic space* (Vol. 639, pp. 65–77). Lecture Notes in Computer Science. Berlin, Heidelberg: Springer. doi:10.1007/3-540-55966-3_3

Couclelis, H., Golledge, R. G., Gale, N., & Tobler, W. (1987). Exploring the anchor-point hypothesis of spatial cognition. *Journal of Environmental Psychology*, *7*(2), 99–122. doi:10.1016/S0272-4944(87)80020-8

Coull, J. T., Charras, P., Donadieu, M., Droit-Volet, S., & Vidal, F. (2015, November). SMA selectively codes the active accumulation of temporal, not spatial, magnitude. *Journal of Cognitive Neuroscience*, *27*(11), 2281–2298. doi:10.1162/jocn_a_00854

Currier, K. E. (2011, December). *Auditory representation in spatial applications* (Master's thesis, University of California, Santa Barbara).

de Campo, A. (2007, February). A data sonification design space map. In A. Hunt & T. Hermann (Eds.), *Proceedings of the 2nd Interactive Sonification Workshop*. York, United Kingdom.

De Felice, F., Renna, F., Attolico, G., & Distante, A. (2007, March). A haptic/acoustic application to allow blind the access to spatial information. In *Second Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems* (pp. 310–315). Tsukuba, Japan. doi:10.1109/WHC.2007.6

De Martino, F., Moerel, M., Ugurbil, K., Formisano, E., & Yacoub, E. (2015). Less noise, more activation: multiband acquisition schemes for auditory functional MRI. *Magnetic Resonance in Medicine*, *74*(2), 462–467. doi:10.1002/mrm.25408

Deatherage, B. H. (1972). Auditory and other sensory forms of information presentation. In H. P. Van Cott & R. G. Kinkade (Eds.), *Human engineering guide to equipment design* (Revised, Chap. 4, pp. 123–160). Washington, D.C.: U.S. Government Printing Office.

Delogu, F., Palmiero, M., Federici, S., Plaisant, C., Zhao, H., & Belardinelli, O. (2010, May). Non-visual exploration of geographic maps: does sonification help? *Disability and Rehabilitation: Assistive Technology*, *5*(3), 164–174. doi:10.3109/17483100903100277

Dent, B. D. (1999). *Cartography: thematic map design* (5th edition). Boston, Massachusetts: The McGreaw-Hill Companies, Inc.

Desmond, J. E. & Glover, G. H. (2002, August). Estimating sample size in functional MRI (fMRI) neuroimaging studies: statistical power analyses. *Journal of Neuroscience Methods*, *118*(2), 115–128. doi:10.1016/S0165-0270(02)00121-8

Diaz-Merced, W. L., Candey, R. M., Brickhouse, N., Schneps, M., Mannone, J. C., Brewster, S., & Kolenberg, K. (2011). Sonification of astronomical data. *Proceedings of the International Astronomical Union*, *7*(S285), 133–136. doi:10.1017/S1743921312000440

Dingler, T., Lindsay, J., & Walker, B. N. (2008, June). Learnabiltiy of sound cues for environmental features: auditory icons, earcons, spearcons, and speech. In *Proceedings of the 14th International Conference on Auditory Display*. Paris, France. Retrieved from http://www.icad.org/node/2343

Donders, F. (1969). On the speed of mental processes (W. G. Koster, Trans.) In *Acta psychologica* (Vol. 30, pp. 412–431). (Original work published 1868). doi:10.1016/0001-6918(69)90065-1

Donkin, K. (1970). Are map libraries obsolete? *Proceedings of the Annual Conference - Association of Canadian Map Libraries*, 60–62.

Drake, C. & Botte, M.-C. (1993, May 1). Tempo sensitivity in auditory sequences: evidence for a multiple-look model. *Perception & Psychophysics*, *54*(3), 277–286. doi:10.3758/BF03205262

Dubus, G. & Bresin, R. (2013, December). A systematic review of mapping strategies for the sonification of physical quantities. *PLoS ONE*, *8*(12). doi:10.1371/journal.pone.0082491

Edsall, R. (2011). Sounds dangerous: emotion, geovisual analytics, and music. In A. Ruas (Ed.), *International Cartographic Conference* (P-105). Paris, France. Retrieved from http://icaci.org/files/documents/ICC_proceedings/ICC2011/Poster%20Presentations%20PDF/POSTERS%20SESSION%202/P-105.pdf

Eklund, A., Nichols, T. E., & Knutsson, H. (2016). Cluster failure: why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences*, *113*(28), 7900–7905. doi:10.1073/pnas.1602413113

Ekstrom, A. (2010). How and when the fMRI BOLD signal relates to underlying neural activity: the danger in dissociation. *Brain Research Reviews*, *62*(2), 233–244. doi:10.1016/j.brainresrev.2009.12.004

Epstein, R., Harris, A., Stanley, D., & Kanwisher, N. (1999). The parahippocampal place area: recognition, navigation, or encoding? *Neuron*, *23*(1), 115–125. doi:10.1016/S0896-6273(00)80758-8

Essinger, R. (1986, September). The philosophy and requirement of computer-aided graphic design in cartography. In M. Blakemore (Ed.), *Proceedings Volume 1: Hardware, Data Capture and Management Techniques* (pp. 189–196). London, England: Auto Carto London Ltd. Retrieved from http://mapcontext.com/autocarto/proceedings/auto-carto-london-vol-1/pdf/the-philosophy-and-requirements-of-computer-aided.pdf

Evers, J. (Ed.) (2011). *Equator*. National Geographic Encyclopedia. Retrieved from https://www.nationalgeographic.org/encyclopedia/equator/

Fabrikant, S. I., Montello, D. R., Ruocco, M., & Middleton, R. S. (2004). The distance–similarity metaphor in network-display spatializations. *Cartography and Geographic Information Science*, *31*(4), 237–252. doi:10.1559/1523040042742402

Fisher, P. F. (1994). Hearing the reliability in classified remotely sensed images. *Cartography and Geographic Information Systems*, *21*(1), 31–36. doi:10.1559/152304094782563975.

Flowers, J. H. (2005, July). Thirteen years of reflection on auditory graphing: promises, pitfalls, and potential new directions. In E. Brazil (Ed.), *Proceedings of the 11th International Conference on Auditory Display (ICAD 2005).* Limerick, Ireland. Retrieved from http://hdl.handle.net/1853/50188

Flowers, J. H., Buhman, D. C., & Turnage, K. D. (1997). Cross-modal equivalence of visual and auditory scatterplots for exploring bivariate data samples. *Human Factors*, *39*(3), 341–351. doi:10.1518/001872097778827151

Frascara, J. & Takach, B. S. (1993). The design of tactile map symbols for visually impaired people. *Information Design Journal*, *7*(1), 67–75. doi:10.1075/idj.7.1.06fra

Frauenberger, C., Putz, V., Hoeldrich, R., & Stockman, T. (2005). Interaction patterns for auditory user interfaces. In E. Brazil (Ed.), *Proceedings of the 11th International Conference on Auditory Display (ICAD2005)* (pp. 154–161). Limerick, Ireland. Retrieved from http://hdl.handle.net/1853/50161

French, A. P. (1982). How far away is the horizon? *American Journal of Physics*, *50*(9), 795–799. doi:10.1119/1.13099

Frye, C. (2015, March). Characterizing regions automatically using global GIS data and natural language generation. *Annual Meeting of the Association of American Geographers.*

Gattis, M. (2001). Reading pictures: constraints on mapping conceptual and spatial schemas. In *Spatial schemas and abstract thought* (Chap. 9). Cambridge, Massachusetts: MIT Press.

Gaver, W. W. (1989, March). The SonicFinder: an interface that uses auditory icons. *Human-Compututer Interaction*, *4*(1), 67–94. doi:10.1207/s15327051hci0401_3

Gersmehl, P. J. (1985). The data, the reader, and the innocent bystander - a parable for map users. *Professional Geographer*, *37*(3), 329–334.

Gersmehl, P. J. (2011). "Wormholes" in the common core: spatial reasoning, literacy, and mathematics education. *Revista Virtual - Geografía, Cultura Y Educación*, *1*(2), 18–53. Retrieved from http://anekumene.com/index.php/revista/article/view/26/25.

Gersmehl, P. J. & Gersmehl, C. A. (2006). Wanted: a concise list of neurologically defensible and assessable spatial thinking skills. *Research in Geographic Education*, *8*, 5–38. Retrieved from http://rge.grosvenor.txstate.edu/Issues/Volume-8.html

Gersmehl, P. J. & Gersmehl, C. A. (2007). Spatial thinking by young children: neurologic evidence for early development and "educability". *Journal of Geography*, *106*(5), 181–191. doi:10.1080/00221340701809108

Glover, G. H. (1999). Deconvolution of impulse response in event-related bold fmri1. *NeuroImage*, *9*(4), 416–429. doi:10.1006/nimg.1998.0419

Goldstein, E. B. (2014). *Sensation and perception* (9th edition). Belmont, California: Wadsworth Cengage Learning.

Golledge, R. G. (2002). The nature of geographic knowledge. *Annals of the Association of American Geographers*, *92*(1), 1–14. doi:10.1111/1467-8306.00276

Golledge, R. G., Marsh, M., & Battersby, S. (2008). Matching geospatial concepts with geographic eduational needs. *Geographical Research*, *46*(1), 85–98. doi:10.1111/j.1745-5871.2007.00494.x

Gorgolewski, K. J., Auer, T., Calhoun, V. D., Craddock, R. C., Das, S., Duff, E. P., ... Poldrack, R. A. (2016, June). The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific Data*, *3*. doi:10.1038/sdata.2016.44

Goudarzi, V. (2016). Exploration of sonification design process through an interdisciplinary workshop. In *Proceedings of the Audio Mostly 2016* (pp. 147–153). Norrköping, Sweden: ACM. doi:10.1145/2986416.2986422

Guo, D., Chen, J., MacEachren, A. M., & Liao, K. (2006, November). A visualization system for space-time and multivariate patterns (VIS-STAMP). *IEEE Transactions on Visualization and Computer Graphics*, *12*(6), 1461–1474. doi:10.1109/TVCG.2006.84

Guo, Q., Thabane, L., Hall, G., McKinnon, M., Goeree, R., & Pullenayegum, E. (2014). A systematic review of the reporting of sample size calculations and corresponding data components in observational functional magnetic resonance imaging studies. *NeuroImage*, *86* (Supplement C), 172–181. doi:10.1016/j.neuroimage.2013.08.012

Hadhazy, A. (2014, April). Heavenly sounds: hearing astronomical data can lead to scientific insights. *Scientific American*. Retrieved from https://www.scientificamerican.com/article/heavenly-sounds-hearing-astronomical-data-can-lead-to-scientific-insights/

Hall, D. A. (2006). fMRI of the auditory cortex. In S. H. Faro & F. B. Mohamed (Eds.), *Functional MRI: basic principles and clinical applications* (Chap. 14, pp. 364–393). New York, New York: Springer Science+Business Media.

Hall, D. A., Haggard, M. P., Akeroyd, M. A., Palmer, A. R., Summerfield, A. Q., Elliott, M. R., … Bowtell, R. W. (1999). "Sparse" temporal sampling in auditory fMRI. *Human Brain Mapping*, *7*, 213–223. Retrieved from https://pdfs.semanticscholar.org/0f68/7945bd663140b603ee3ad7dbee41f659308e.pdf

Hartley, T., Lever, C., Burgess, N., & O'Keefe, J. (2013). Space in the brain: how the hippocampal formation supports spatial cognition. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *369*(1635). doi:10.1098/rstb.2012.0510

Hegarty, M., Montello, D. R., Richardson, A. E., Ishikawa, T., & Lovelace, K. (2006). Spatial abilities at different scales: individual differences in aptitude-test performance and spatial-layout learning. *Intelligence*, *34*(2), 151–176. doi:10.1016/j.intell.2005.09.005

Hegg, J. C., Middleton, J., Robertson, B. L., & Kennedy, B. P. (2018). The sound of migration: exploring data sonification as a means of interpreting multivariate salmon movement datasets. *Heliyon*, *4*(2), e00532. doi:10.1016/j.heliyon.2018.e00532

Hermann, T., Hunt, A., & Neuhoff, J. G. (Eds.). (2011). *The sonification handbook*. New York, New York: Academic Press. Retrieved from http://sonification.de/handbook/

Jenkinson, M. (2002). Fast, automated, N-dimensional phase-unwrapping algorithm. *Magnetic Resonance in Medicine*, *49*(1), 193–197. doi:10.1002/mrm.10354

Jenkinson, M. (2004). Improving the registration of B0-disorted EPI images using calculated cost function weights. *Tenth International Conference on Functional Mapping of the Human Brain*.

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, *17*(2), 825–841. doi:10.1006/nimg.2002.1132

Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W., & Smith, S. M. (2012, August). FSL. *NeuroImage*, *62*(2), 782–790. doi:10.1016/j.neuroimage.2011.09.015

Judd, C. M., McClelland, G. H., & Ryan, C. S. (2009). *Data analysis: a model comparison approach* (2nd edition). New York, New York: Routledge.

Kanwisher, N. (2010). Functional specificity in the human brain: a window into the functional architecture of the mind. *Proceedings of the National Academy of Sciences*, *107*(25), 11163–11170. doi:10.1073/pnas.1005062107

Keenan, J. P., Thangaraj, V., Halpern, A. R., & Schlaug, G. (2001). Absolute pitch and planum temporale. *NeuroImage*, *14*(6), 1402–1408. doi:10.1006/nimg.2001.0925

Keller, J. M., Buck, A. R., Zare, A., & Popescu, M. (2014, December). A human geospatial predictive analytics framework with application to finding medically underserved areas. In *IEEE Symposium on Computational Intelligence in Big Data (CIBD)* (pp. 1–6). doi:10.1109/CIBD.2014.7011525

Kitchin, R. & Dodge, M. (2007). Rethinking maps. *Progress in Human Geography*, *31*(3), 331–344. doi:10.1177/0309132507077082

Koláčný, A. (1969). Cartographic information—a fundamental concept and term in modern cartography. *The Cartographic Journal*, *6*(1), 47–49. doi:10.1179/caj.1969.6.1.47

Kramer, G., Walker, B. N., Bonebright, T., Cook, P., Flowers, J. H., Miner, N., & Neuhoff, J. (1997). Sonification report: status of the field and research agenda. Retrieved from http://www.icad.org/websiteV2.0/References/nsf.html

Krygier, J. B. (1994). Sound and geographic visualization. In A. M. MacEachren & D. R. Taylor (Eds.), *Visualization in modern cartography* (Chap. 8, pp. 149–166). Oxford, United Kingdom: Pergamon.

Kulhavy, R. W. & Stock, W. A. (1996). How cognitive maps are learned and remembered. *Annals of the Association of American Geographers*, *86*(1), 123–145. doi:10.1111/j.1467-8306.1996.tb01748.x

Langers, D. R. M., Van Dijk, P., & Backes, W. H. (2005, January). Interactions between hemodynamic responses to scanner acoustic noise and auditory stimuli in functional magnetic resonance imaging. *Magnetic Resonance in Medicine*, *53*(1), 49–60. doi:10.1002/mrm.20315

Larkman, D. J., Hajnal, J. V., Herlihy, A. H., Coutts, G. A., Young, I. R., & Ehnholm, G. (2001, January). Use of multicoil arrays for separation of signal from multiple slices simultaneously excited. *Journal of Magnetic Resonance Imaging*, *13*(2), 313–317. doi:10.1002/1522-2586(200102)13:2⟨313::AID-JMRI1045⟩3.0.CO;2-W

Levitin, D. J. (2002). Memory for musical attributes. In D. J. Levitin (Ed.), *Foundations in cognitive psychology* (Chap. 13, pp. 295–310). Cambridge, Massachusetts: MIT Press.

Liem, F., Lutz, K., Luechinger, R., Jäncke, L., & Meyer, M. (2012, April). Reducing the interval between volume acquisitions improves "sparse" scanning protocols in event-related auditory fMRI. *Brain Topography*, *25*(2), 182–193. doi:10.1007/s10548-011-0206-x

Liu, T. T., Frank, L. R., Wong, E. C., & Buxton, R. B. (2001). Detection power, estimation efficiency, and predictability in event-related fmri. *NeuroImage*, *13*(4), 759–773. doi:10.1006/nimg.2000.0728

Lobben, A. (2003). Classification and application of cartographic animation. *The Professional Geographer*, *55*(3), 318–328. doi:10.1111/0033-0124.5503016

Lobben, A., Brittell, M., & Perdue, N. (2015). Inclusive cartographic design: overcoming ocular-centric cartographies. In C. Robbi Sluter, B. C. Madureira Cruz, & M. P. Leal de Menezes (Eds.), *Cartography - Maps Connecting the World: Proceedings of the 27th International Cartographic Conference* (Chap. 7, pp. 89–98). Springer International Publishing. doi:10.1007/978-3-319-17738-0_7

Lobben, A., Lawrence, M., & Olson, J. M. (2009). fMRI and human subjects research in cartography. *Cartographica: The International Journal for Geographic Information and Geovisualization*, *44*(3), 159–169. doi:10.3138/carto.44.3.159

Lobben, A., Lawrence, M., & Pickett, R. (2014). The map effect. *Annals of the Association of American Geographers*, *104*(1), 96–113. doi:10.1080/00045608.2013.846172

Lobben, A., Olson, J. M., & Huang, J. (2005). Using fMRI in cartographic research. In *Proceedings of the 22nd International Cartographic Conference*. A Coruña, Spain. Retrieved from http://www.cartesia.es/geodoc/icc2005/pdf/oral/TEMA1/SESSION1/AMY%20LOBBEN1.pdf

MacEachren, A. M. (1982a). Map complexity: comparison and measurement. *The American Cartographer*, *9*(1), 31–46. Retrieved from http://www.geovista.psu.edu/publications/MacEachren/MacEachren_MapComplexity_82.pdf

MacEachren, A. M. (1982b). The role of complexity and symbolization method in thematic map effectiveness. *Annals of the Association of American Geographers*, *72*(4), 495–513. doi:10.1111/j.1467-8306.1982.tb01841.x

MacEachren, A. M. (1994). Cartographic language. In *Some truth with maps: a primer on symbolization and design* (Chap. 2). Washington, D.C.: Association of American Geographers.

MacEachren, A. M. (1995). *How maps work: representation, visualization, and design*. New York, New York: Guilford Press.

MacEachren, A. M., Wachowicz, M., Edsall, R., Haug, D., & Masters, R. (1999). Constructing knowledge from multivariate spatio-temporal data: integrating geographical visualization with knowledge discovery in database methods. *International Journal of Geographical Information Science*, *13*(4), 311–334. doi:10.1080/136588199241229

Maus, B. & van Breukelen, G. J. P. (2013). Optimal design for functional magnetic resonance imaging experiments: methodology, challenges, and future perspectives. *Zeitschrift für Psychologie*, *221*(3), 174–189. doi:10.1027/2151-2604/a000145

McCarthy, P. (2018, October). FSLeyes. Zendo. doi:10.5281/zenodo.1470762

McJury, M. & Shellock, F. G. (2000). Auditory noise associated with MR procedures: a review. *Journal of Magnetic Resonance Imaging*, *12*(1), 37–45. doi:10.1002/1522-2586(200007)12:1⟨37::AID-JMRI5⟩3.0.CO;2-I

Medyckyj-Scott, D. J. & Blades, M. (1991). Cognitive representations of space in the design and use of geographical information systems. In D. Diaper & N. Hammond (Eds.), *People and Computers VI: Proceedings of HCI '91* (pp. 421–433). Cambridge, Massachusetts: Cambridge University Press.

Meijer, P. B. L. (1992, February). An experimental system for auditory image representations. *IEEE Transactions on Biomedical Engineering*, *39*(2), 112–121.

Moelker, A. & Pattynama, P. M. (2003). Acoustic noise concerns in functional magnetic resonance imaging. *Human Brain Mapping*, *20*(3), 123–141. doi:10.1002/hbm.10134

Montello, D. R. (1997). The perception and cognition of environmental distance: direct sources of information. In S. Hirtle & A. Frank (Eds.), *Spatial information theory a theoretical basis for GIS* (Vol. 1329, pp. 297–311). Lecture Notes in Computer Science. Springer Berlin / Heidelberg. doi:10.1007/3-540-63623-4_60

Montello, D. R. (2002). Cognitive map-design research in the twentieth century: theoretical and empirical approaches. *Cartography and Geographic Information Science*, *29*(3), 283–304. Retrieved from http://intraspec.ca/12montello.pdf

Montello, D. R. (2016). Behavioral methods for spatial cognition research. In R. Gifford (Ed.), *Research methods for environmental psychology* (Chap. 9, pp. 161–172). Hoboken, New Jersey: Wiley.

Montello, D. R. & Raubal, M. (2012). Functions and applications of spatial cognition. In D. Waller & L. Nadel (Eds.), *Handbook of spatial cognition* (Chap. 14). American Psychological Association.

Morrison, J. L. (1974). Changing philosophical-technical aspects of thematic cartography. *The American Cartographer*, *1*(1), 5–14. doi:10.1559/152304074784107854

Mowrer, E. R. (1938). The isometric map as a technique of social research. *American Journal of Sociology*, *44*(1), 86–96. Retrieved from http://www.jstor.org/stable/2768124

Muehrcke, P. (1973). Visual pattern comparison in map reading. *Proceedings of the Association of American Geographers*, *5*, 190–194.

Mumford, J. A. (2010). First-level fMRI modeling. *UCLA Advanced NeuroImaging Summer School*. Retrieved from http://mumford.fmripower.org/2010_summercourse/lev1_2010.pdf

Mynatt, E. D. & Edwards, W. K. (1992). Mapping GUIs to auditory interfaces. In *Proceedings of the 5th Annual ACM Symposium on User Interface Software and Technology* (pp. 61–70). Monteray, California: ACM. doi:10.1145/142621.142629

Nees, M. A. & Walker, B. N. (2008, June). Encoding and representation of information in auditory graphs: descriptive reports of listener strategies for understanding data. In *Proceedings of the 14th International Conference on Auditory Display*. Paris, France. Retrieved from http://hdl.handle.net/1853/49907

Nees, M. A. & Walker, B. N. (2009). Auditory interfaces and sonification. *The Universal access handbook*, 507–521. Retrieved from http : / / sonify . psych . gatech . edu / ~walkerb / publications / pdfs / NeesWalker - UniversalAccessChapter2007-submitted.pdf

Nees, M. A. & Walker, B. N. (2011, July). Mental scanning of sonifications reveals flexible encoding of nonspeech sounds and a universal per-item scanning cost. *Acta Psychologica*, *137*(3), 309–317. doi:10.1016/j.actpsy.2011.03.012

Nettl, B. (1960, October). Musical cartography and the distribution of music. *Southwestern Journal of Anthropology*, *16*(3), 338–347. doi:10.2307/3629036

Newcombe, N. S. (1985). Methods for the study of spatial cognition. In R. Cohen (Ed.), *The development of spatial cognition* (Chap. 11, pp. 277–300). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Nielsen, J., Clemmensen, T., & Yssing, C. (2002). Getting access to what goes on in people's heads?: reflections on the think-aloud technique. In *Proceedings of the second Nordic conference on human-computer interaction* (pp. 101–110). NordiCHI '02. Aarhus, Denmark: ACM. doi:10.1145/572020.572033

Nystuen, J. D. (1968). Identification of some fundamental spatial concepts. In B. J. L. Berry & D. F. Marble (Eds.), *Spatial analysis: a reader in statistical geography* (Chap. 3, pp. 35–41). Englewood Cliffs, New Jersey: Prentice Hall.

Ogao, P. J. & Kraak, M.-J. (2001). Applying a creative cognition approach to exploratory cartography tool development. In *Proceedings of the 20th International Cartographic Conference*. Beijing, China. Retrieved from http://icaci.org/files/documents/ICC_proceedings/ICC2001/icc2001/file/f16010.pdf

Ogawa, S., Lee, T. M., Kay, A. R., & Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences*, *87*(24), 9868–9872.

Ojala, S., Lahtinen, R., & Hirn, H. (2016). Tactile maps – safety and usability, 15–22. doi:10.1007/978-3-319-44672-1_2

O'Keefe, J. (2014, December). Spatial cells in the hippocampal formation. *Nobel Lecture*. Retrieved from https://www.nobelprize.org/nobel_prizes/medicine/laureates/2014/okeefe-lecture.pdf

O'Keefe, J. & Dostrovsky, J. (1971). The hippocampus as a spatial map. preliminary evidence from unit activity in the freely-moving rat. *Brain Research*, *34*(1), 171–175. doi:10.1016/0006-8993(71)90358-1

Olson, J. M. (1975, June). Autocorrelation and visual map complexity. *Annals of the Association of American Geographers*, *65*(2), 189–204. Retrieved from http://www.jstor.org/stable/2562081

Pasqualotto, A., Spiller, M. J., Jansari, A. S., & Proulx, M. J. (2013). Visual experience facilitates allocentric spatial representation. *Behavioural Brain Research*, *236*, 175–179. doi:10.1016/j.bbr.2012.08.042

Paté, A., Boschi, L., Carrou, J.-L. L., & Holtzman, B. (2016). Categorization of seismic sources by auditory display: a blind test. *International Journal of Human-Computer Studies, Special issue on Data Sonification and Sound Design in Interactive Systems*, *85*, 57–67. doi:10.1016/j.ijhcs.2015.08.002

Peirce, J. W. (2007). PsychoPy–psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1–2), 8–13. doi:10.1016/j.jneumeth.2006.11.017

Peirce, J. W. (2009, January). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, *2*, 10. doi:10.3389/neuro.11.010.2008

Perrachione, T. & Ghosh, S. (2013). Optimized design and analysis of sparse-sampling fMRI experiments. *Frontiers in Neuroscience*, *7*, 55. doi:10.3389/fnins.2013.00055

Peuquet, D. J. (1984). A conceptual framework and comparison of spatial data models. *Cartographica: The International Journal for Geographic Information and Geovisualization*, *21*(4), 66–113. doi:10.3138/D794-N214-221R-23R5

Peuquet, D. J. (2001). Making space for time: issues in space-time data representations. *GeoInformatica*, *5*(1), 11–32. doi:10.1023/A:1011455820644

Piaget, J. & Inhelder, B. (1956). *The Child's Conception of Space* (F. J. Langdon and J. L. Lunzer, Trans.) (Chap. 1–2). London, England: Routledge & Kegan Paul. (Original work published 1948)

Pollack, I. & Ficks, L. (1954). Information of elementary multidimensional auditory displays. *The Journal of the Acoustical Society of America*, *26*(2), 155–158. doi:10.1121/1.1907300

Rice, M., Jacobson, R. D., Golledge, R. G., & Jones, D. (2005). Design considerations for haptic and auditory map interfaces. *Cartography and Geographic Information Science*, *32*(4), 381–391. doi:10.1559/152304005775194656

Richter, K.-F. & Winter, S. (2014). Cognitive aspects: how people perceive, memorize, think and talk about landmarks. In *Landmarks* (pp. 41–108). Springer International Publishing. doi:10.1007/978-3-319-05732-3_3

Rigas, D. & Alty, J. L. (2005). The rising pitch metaphor. *International Journal of Human-Computer Studies*, *62*(1), 1–20. doi:10.1016/j.ijhcs.2004.06.004

Rinck, M. (2005). Spatial situation models. In P. Shah & A. Miyake (Eds.), *The Cambridge handbook of visuospatial thinking*. New York, New York: Cambridge University Press.

Robare, P. & Forlizzi, J. (2009, January). Sound in computing: a short history. *interactions*. Timelines, *16*(1), 62–65. doi:10.1145/1456202.1456218

Robinson, A. H. (1975, September). Map design. In J. C. Kavaliunas (Ed.), *Proceedings of the International Symposium on Computer-Assisted Cartography: AutoCarto II* (pp. 9–14). U.S. Department of Commerce, Bureau of the Census.

Robinson, A. H. (1986). The look of maps: an examination of cartographic design. *The American Cartographer*, *13*(3), 280–280. doi:10 . 1559 / 152304086783899881

Robinson, A. H. & Bryson, R. A. (1957). A method for describing quantitatively the correspondence of geographical distributions. *Annals of the Association of American Geographers*, *47*(4), 379–391.

Robinson, A. H., Morrison, J. L., Muehrcke, P. C., Kimerling, A. J., & Guptill, S. C. (1995). *Elements of cartography* (6th edition). Hoboken, New Jersey: John Wiley & Sons, Inc.

Rozovskaya, R. I. & Pechenkova, E. V. (2012). Activation of parahippocampal place area when looking at geographic maps: an fMRI study. In *Perception, European Conference on Visual Perception (ECVP)* (Vol. 41, pp. 197–198). doi:10.1177/03010066120410S101

Saltré, F., Duputié, A., Gaucherel, C., & Chuine, I. (2015). How climate, migration ability and habitat fragmentation affect the projected future distribution of European beech. *Global Change Biology*, *21*(2), 897–910. doi:10.1111/gcb.12771

Samuels, D. W., Meintjes, L., Ochoa, A. M., & Porcello, T. (2010). Soundscapes: toward a sounded anthropology. *Annual Review of Anthropology*, *39*(1), 329–345. doi:10.1146/annurev-anthro-022510-132230

Savidis, A. & Stephanidis, C. (1995). Developing dual user interfaces for integrating blind and sighted users: the HOMER UIMS. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 106–113). CHI '95. Denver, Colorado: ACM Press/Addison-Wesley Publishing Co. doi:10.1145/223904.223918

Schito, J. & Fabrikant, S. I. (2018). Exploring maps by sounds: using parameter mapping sonification to make digital elevation models audible. *International Journal of Geographical Information Science*, 1–34. doi:10.1080/13658816.2017.1420192

Schneiderman, B. & Plaisant, C. (2010). *Designing the user interface: strategies for effective human-computer interaction* (5th edition). Boston, Massachusetts: Addison-Wesley.

Schwarzbauer, C., Davis, M. H., Rodd, J. M., & Johnsrude, I. (2006, February). Interleaved silent steady state (ISSS) imaging: a new sparse imaging method applied to auditory fMRI. *NeuroImage*, *29*(3), 774–782. doi:10.1016/j.neuroimage.2005.08.025

Shellock, F. G. & Crues, J. V. (2004). MR procedures: biologic effects, safety, and patient care. *Radiology*, *232*, 635–652. doi:10.1148/radiol.2323030830

Sheskin, D. J. (1997). *Handbook of parametric and nonparametric statistical procedures*. Boca Raton, Florida: CRC Press.

Sigsworth, G. (2018, May). On making music with machines. Thelma Schnitzer Hall, Frohnmayer Music Building, University of Oregon. Retrieved from https://calendar.uoregon.edu/event/public_lecture_guy_sigsworth_producer_and_composer_2278

Silson, E. H., Chan, A. W.-Y., Reynolds, R. C., Kravitz, D. J., & Baker, C. I. (2015). A retinotopic basis for the division of high-level scene processing between lateral and ventral human occipitotemporal cortex. *Journal of Neuroscience*, *35*(34), 11921–11935. doi:10.1523/JNEUROSCI.0137-15.2015

Skupin, A. & Fabrikant, S. I. (2003, April). Spatialization methods: a cartographic research agenda for non-geographic information visualization. *Cartography and Geographic Information Science*, *30*(2), 99–119. doi:10.1559/152304003100011081

Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, *17*(3), 143–155. doi:10.1002/hbm.10062

Smith, S. M. & Brady, J. M. (1997). SUSAN—a new approach to low level image processing. *International Journal of Computer Vision*, *23*(1), 45–78. doi:10.1023/A:1007963824710

St. George, S., Crawford, D., Reubold, T., & Giorgi, E. (2017). Making climate data sing: using music-like sonifications to convey a key climate record. *Bulletin of the American Meteorological Society*, *98*(1), 23–27. doi:10.1175/BAMS-D-15-00223.1

Sumikawa, D. A. (1985). *Guidelines for the integration of audio cues into computer user interfaces* (Master's thesis, University of California, Davis).

Supper, A. (2014). Sublime frequencies: the construction of sublime listening experiences in the sonification of scientific data. *Social Studies of Science*, *44*(1), 34–58. doi:10.1177/0306312713496875

Supper, A. (2015). Sound information: sonification in the age of complex data and digital audio. *Information & Culture*, *50*(4), 441–464. doi:10.7560/IC50401

Talavage, T. M., Gonzalez-Castillo, J., & Scott, S. K. (2014, January). Auditory neuroimaging with fMRI and PET. In D. A. Hall & D. R. M. Langers (Eds.), *Hearing Research, special issue: Human auditory neuroimaging* (Vol. 307, pp. 4–15). doi:10.1016/j.heares.2013.09.009

Tatham, A. (1995). Cartography crossing sensory borders. In *Proceedings of the 17th International Cartographic Conference*. Barcelona, Spain. Retrieved from https://icaci.org/files/documents/ICC_proceedings/ICC1995/PDF/Cap256.pdf

Thorndyke, P. W. & Stasz, C. (1980). Individual differences in procedures for knowledge acquisition from maps. *Cognitive Psychology*, *12*(1), 137–175. doi:http://dx.doi.org/10.1016/0010-0285(80)90006-7

Timpf, S. & Frank, A. U. (1997). Using hierarchical spatial data structures for hierarchical spatial reasoning. In S. C. Hirtle & A. U. Frank (Eds.), *Spatial information theory a theoretical basis for GIS* (Vol. 1329, pp. 69–83). COSIT 1997. Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/3-540-63623-4_43

Tobler, W. R. (1959). Automation and cartography. *Geographical Review*, *49*(4), 526–534. Retrieved from http://www.jstor.org/stable/212211

Todd, N., Moeller, S., Auerbach, E. J., Yacoub, E., Flandin, G., & Weiskopf, N. (2016, January). Evaluation of 2D multiband EPI imaging for high-resolution, whole-brain, task-based fMRI studies at 3T: sensitivity and slice leakage artifacts. *NeuroImage*, *124*(A), 32–42. doi:10.1016/j.neuroimage.2015.08.056

Tomasi, D., Caparelli, E. C., Chang, L., & Ernst, T. (2005, August). fMRI-acoustic noise alters brain activation during working memory tasks. *NeuroImage*, *27*(2), 377–386. doi:10.1016/j.neuroimage.2005.04.010

Turnbull, W. W. (1944). Pitch discrimination as a function of tonal duration. *Journal of Experimental Psychology*, *34*(4), 302–316. doi:10.1037/h0063434

Tversky, B. (1993). Cognitive maps, cognitive collages, and spatial mental models. In A. Frank & I. Campari (Eds.), *Spatial information theory a theoretical basis for GIS* (Vol. 716, pp. 14–24). Lecture Notes in Computer Science. Springer Berlin Heidelberg. doi:10.1007/3-540-57207-4_2

Ulmer, J. L., Biswal, B. B., Mark, L. P., Mathews, V. P., Prost, R. W., Millen, S. J., … Horzewski, D. (1998, May). Acoustic echoplanar scanner noise and pure tone hearing thresholds: the effects of sequence repetition times and acoustic noise rates. *Journal of Computer Assisted Tomography*, *22*(3), 480–486. Retrieved from https://journals.lww.com/jcat/Abstract/1998/05000/Acoustic_Echoplanar_Scanner_Noise_and_Pure_Tone.22.aspx

Vasiliev, I. (1996). Design issues to be considered when mapping time. In *Design issues to be considered when mapping time* (Chap. 11, pp. 137–146). Chichester, New York: Wiley.

Visscher, K. M., Miezin, F. M., Kelly, J. E., Buckner, R. L., Donaldson, D. I., McAvoy, M. P., … Petersen, S. E. (2003). Mixed blocked/event-related designs separate transient and sustained activity in fMRI. *NeuroImage*, *19*, 1694–1708. doi:10.1016/S1053-8119(03)00178-2

Walker, B. N. & Kramer, G. (2006, October). Mappings and metaphors in auditory displays: an experimental assessment. *ACM Transactions on Applied Perception*, *2*(4), 407–412.

Walker, B. N. & Mauney, L. M. (2010, March). Universal design of auditory graphs: a comparison of sonification mappings for visually impaired and sighted listeners. *ACM Transactions on Accessible Computing*, *2*(3), 12:1–12:16. doi:10.1145/1714458.1714459

Wickham, H. (2014, August). Tidy data. *Journal of Statistical Software*, *59*(10). Retrieved from https://www.jstatsoft.org/article/view/v059i10

Wojtczak, M., Mehta, A. H., & Oxenham, A. J. (2017). Rhythm judgments reveal a frequency asymmetry in the perception and neural coding of sound synchrony. *Proceedings of the National Academy of Sciences*, *114*(5), 1201–1206. doi:10.1073/pnas.1615669114

Woo, C.-W., Krishnan, A., & Wager, T. D. (2014). Cluster-extent based thresholding in fmri analyses: pitfalls and recommendations. *NeuroImage*, *91*, 412–419. doi:10.1016/j.neuroimage.2013.12.058

Wood, M. (1968). Visual perception and map design. *Cartographic Journal, The*, *5*(1), 54–64.

Wood, M. (1972). Human factors in cartographic communication. *The Cartographic Journal*, *9*(2), 123–132.

Woolrich, M. W., Ripley, B. D., Brady, M., & Smith, S. M. (2001). Temporal autocorrelation in univariate linear modeling of fmri data. *NeuroImage*, *14*(6), 1370–1386. doi:10.1006/nimg.2001.0931

Woolrich, M., Behrens, T., Beckmann, C., Jenkinson, M., & Smith, S. (2004, April). Multi-level linear modelling for FMRI group analysis using Bayesian inference. *NeuroImage*, *21*(4), 1732–1747. PMID 15050594. doi:10.1016/j.neuroimage.2003.12.023

Worsley, K. (2001). Statistical analysis of activation images. In P. Jezzard, P. Matthews, & S. Smith (Eds.), *Functional MRI: an introduction to methods* (Chap. 14). Oxford University Press.

Yeung, E. S. (1980). Pattern recognition by audio representation of multivariate analytical data. *Analytical Chemistry*, *52*(7), 1120–1123. doi:10.1021/ac50057a028

Zhang, H., Copara, M., & Ekstrom, A. D. (2012, September). Differential recruitment of brain networks following route and cartographic map learning of spatial environments. *PLoS ONE*, *7*(9), e44886. doi:10.1371/journal.pone.0044886

Zhao, H., Plaisant, C., Schneiderman, B., & Duraiswami, R. (2004, July). Sonification of geo-referenced data for auditory information seeking: design principle and pilot study. In S. Barrass & P. Vickers (Eds.), *Proceedings of the 10th International Conference on Auditory Display*. Sydney, Australia. Retrieved from http://hdl.handle.net/1853/50918

Zhao, H., Plaisant, C., Shneiderman, B., & Lazar, J. (2008, May). Data sonification for users with visual impairment: a case study with georeferenced data. *ACM Transactions on Computer-Human Interaction*, *15*(1), 4:1–4:28. doi:10.1145/1352782.1352786

Zhao, H., Smith, B. K., Norman, K., Plaisant, C., & Schneiderman, B. (2004). Listening to maps: user evaluation of interactive sonifications of geo-referenced data. *IEEE Transactions on Media*. Institute for Systems Research, University of Maryland (TR 2005-56).