

USABILITY AND USER EXPECTATIONS FOR ONLINE MAP-BASED DATA
VISUALIZATION TOOLS: THE OCEAN NAVIGATOR CASE

by

Alison Faye McLaughlin

Submitted in partial fulfilment of the requirements
for the degree of Master of Information

at

Dalhousie University
Halifax, Nova Scotia
October 2020

This thesis is dedicated to Connie Kavanaugh, my kindergarten teacher, who inspired and nurtured my lifelong love of learning and books.

TABLE OF CONTENTS

LIST OF TABLES	v
LIST OF FIGURES	vi
ABSTRACT	vii
LIST OF ABBREVIATIONS USED	viii
ACKNOWLEDGEMENTS	ix
CHAPTER 1 INTRODUCTION	1
1.1 WEB-BASED MAPPING APPLICATIONS	1
1.2 OVERVIEW	4
CHAPTER 2 LITERATURE REVIEW	6
2.1 GENERAL USABILITY GUIDELINES	6
2.2 MAP APPLICATION USABILITY GUIDELINES.....	12
2.3 USABILITY EVALUATION.....	14
CHAPTER 3 METHODOLOGY	18
3.1 PARTICIPANTS	20
3.2 MATERIALS	21
3.3 PROCEDURE	21
3.4 DATA CLEANING.....	24
CHAPTER 4 RESULTS	26
4.1 DOES OCEAN NAVIGATOR MEET THE TYPICAL NEEDS OF THE USER?	26
4.2 USERS' PERCEPTION OF OCEAN NAVIGATOR AS USABLE.....	32
4.3 SOME USERS TRANSLATE THEIR EXPECTATIONS FROM CONSUMER TOOLS TO NICHE TOOLS LIKE OCEAN NAVIGATOR	36
CHAPTER 5 DISCUSSION AND CONCLUSIONS	39
5.1 FINDINGS	39
5.1.1 Does Ocean Navigator Meet the Typical Needs of the User?	39
5.1.2 Users' Perception of Ocean Navigator as Usable	43
5.1.3 Some Users Translate Their Expectations from Consumer Tools to Niche Tools Like Ocean Navigator	45
5.2 LIMITATIONS.....	46
5.3 SUMMARY OF RECOMMENDATIONS	48
5.4 SUMMARY OF FUTURE RESEARCH SUGGESTIONS.....	49

REFERENCES	51
APPENDIX A Test Script	56
APPENDIX B Consent Forms.....	58
CONSENT FORM – SECOND STREAM	58
CONSENT FORM – FIRST STREAM	60
APPENDIX C Recruitment	64
INDIRECT RECRUITMENT – FIRST STREAM.....	64
DIRECT RECRUITMENT – FIRST STREAM.....	65
MEOPEER RECRUITMENT EMAIL – FIRST STREAM	66
MEOPEER RECRUITMENT FACEBOOK POST – FIRST STREAM	66
APPENDIX D Task Descriptions.....	68
APPENDIX E System Usability Scale	69
APPENDIX F Initial Codes	71
APPENDIX G Codes Discovered During Phase One of Coding	72
APPENDIX H Renaming of Codes for Phase Two.....	73
APPENDIX I Codes Used for Phase Two	75
APPENDIX J Final List of Codes with Definitions	77

LIST OF TABLES

Table 1:	Time (ratio) participants took to complete tasks relative to an expert user.....	28
Table 2:	Participant's SUS score and average completion time relative to an expert user.....	32

LIST OF FIGURES

Figure 1	A comparison of the adjective ratings, acceptability scores, and school grading scales, in relation to the average SUS score.....	16
Figure 2	Screenshot of the date selection widget from worldview.earthdata.nasa.gov	38

ABSTRACT

In the past, mapping applications and programs were primarily used by professionals in the geographical information world to assist in the type of geospatial analytics research they would be performing. However, mapping applications are becoming more accessible to and more utilized by the general public. Ocean Navigator is a modern mapping application that helps visualize various types of ocean data. This study assesses the usability of Ocean Navigator through a think aloud participant observation method paired with the well-established Likert scale tool, the System Usability Scale. It was found that participants are able to meet the typical information needs of the expected user, users perceive Ocean Navigator as usable but not extremely so, and some users translate their expectations from consumer tools to niche tools like Ocean Navigator. These findings are contextualized with the previous literature, and limitations are discussed. Recommendations for immediate modifications and future research are presented.

LIST OF ABBREVIATIONS USED

SUS	System Usability Scale
UCD	User-centered design

ACKNOWLEDGEMENTS

Mike Smit (my co-supervisor): thank you for taking me on as a thesis student and believing in me when I didn't believe in myself.

Jennifer Grek Martin (my co-supervisor): thank you for being a source of consistency, positivity, and guidance.

Sandra Toze (member of my supervisory committee): thank you for supporting me in so many ways.

Sonja McLaughlin (my mom): thank you for helping in every way you could think of. Whether it was calling to make sure I was awake and working, proof reading something, helping me brainstorm, or so many other things, you have been a constant support through my entire degree (and life).

Steven McDonald and Elphege Bernard-Wesson (fellow thesis-writers): thank you for going through this with me. You both encouraged me when it felt like this was impossible, spent countless hours writing with me, and celebrated the small victories along the way. We did it guys, and I am so proud of us!

Samuel Bodapati (my fiancé): thank you for loving and supporting me through the worst of this. You've taken on more than your fair share of household tasks so that I could write for hours and hours and gone with me to the gym, to get food, or wherever I needed to go to take a break and get away from things.

1.1 WEB-BASED MAPPING APPLICATIONS

Given today's prevalent use of technology, web-based maps and map applications are ubiquitous in western society. With over 5,000,000,000 installs from the Google Play store alone (Maps - Navigate & Explore, 2019), Google Maps is never further away than our pockets or handbags. However, in the past, mapping applications and programs were primarily used by professionals in the geographical information world to assist in the type of geospatial analytics research they would be performing.

Until recently the process of accurate map-making was long, complex and expensive. It required considerable skill in the use of photogrammetric techniques, expensive equipment for observation and analysis and substantial investment in large-format printing. Today, however, almost all of these constraints have been removed (Goodchild, 2009, p. 86).

More and more, these types of programs are being used by the general public, or by researchers whose area of expertise is not geographical information systems. Google Maps is a very broad tool that allows many basic functions to be performed but tools exist now that are more specific. These specific tools allow for more in depth analysis and discovery.

Ocean Navigator (<http://navigator.oceansdata.ca/>) is an open-source data map tool that aims to be a “modern, user-friendly, interactive online map, displaying scientific ocean data gathered from various models around the world” (Ocean Navigator, n.d.). Ocean models are

theoretical representations of the ocean, using mathematics to understand the ocean and understanding how that mathematics relates to the real world. Ocean Navigator is designed to enable the discovery and accessibility of over 40 terabytes of ocean data from ocean prediction systems (Ocean Navigator, n.d.). Ocean prediction systems use ocean models to produce forecasts and/or hindcasts of ocean behavior. Ocean Navigator can display several different variables from several different datasets; for example, the Global Ice Ocean Prediction System (GIOPS) daily dataset can provide ice concentration, ice volume, meridional wind, salinity, sea surface height (free surface), sea water velocity, sea water east velocity, sea water north velocity, sea water X velocity, sea water Y velocity, water temperature, wind, and zonal wind (Ocean Navigator, n.d.).

Although the tool was developed by scientists, it is designed for any user, from expert researchers, to laypersons interested in oceans. Ocean Navigator is designed to be simpler and less expensive than traditional map analytics tools like ArcGIS. This is important because it allows scientists who are not specialists in geography to perform analysis they previously would have needed to consult geography experts to perform.

Ocean Navigator is unique because it allows you to visualize four-dimensional data. Four-dimensional data in this context refers to not only representing X (east-west) and Y (north-south) values, but also representing depth and time. It does this through various visualizations such as diagrams, plots, and profiles.

Ocean Navigator was created by experts at Fisheries and Oceans Canada (Ocean Navigator, n.d.) who are likely to have more expertise than the intended users of the application. This means that it is especially important to think about the users' needs since they will likely be different than those of the creators. Therein lies the significance of user-centered design (UCD), which involves focusing the design and development process on the users' needs by involving them along the way (Stone et al., 2005). UCD adds value to applications in many ways, which were summarized by Kramers (2008) as follows:

1. A reduction in the effect of poor and inaccurate assumptions – informed decision making
2. Balance of business and user requirements
3. Developers do not have to evaluate their own designs
4. An increase in user satisfaction and product effectiveness
5. The right product is produced, for the right reasons, for the right users (p. 106)

Ideally, good UCD involves evaluation with users at all levels of design and development (Stone et al., 2005). Users should be involved early in the design process to help determine the requirements for the system, during prototyping to test designs and options, before release of the system to identify minor required tweaks, after the release to identify problems to be addressed in the next version of the system, etcetera (Stone et al., 2005). However, lack of resources and knowledge regarding how to implement this are barriers impacting this involvement (Nivala, Sarjakoski & Sarjakoski, 2007). Integrating user feedback into web mapping application design is very important (Newman et al., 2010) and users say that they would like to give more feedback (Roth, 2015).

It is important to keep in mind that participants may have preconceived ideas from working with popular consumer tools that have some aspects that are similar to Ocean Navigator. For example, participants' expectations for Ocean Navigator could be affected by their ideas of how a mapping platform "should" look and work. As such, usability testing is integral to addressing these preconceived notions about a system.

Usability testing is a method of evaluating how users interact with the interface of a digital environment to identify issues users face in completing specific tasks with ease. It is also intended to identify areas for improvement by testing how satisfied the user is with navigating the digital environment. Identifying whether a tool is easy to use or not is something usability testing can help with. Since the goal of Ocean Navigator is to provide access to data, that data needs to be accessible with ease. Both the ability to carry out tasks and the users' perceptions of the ease with which they carry out those tasks will affect how readily users are willing to subsequently use this tool. Since such tools are designed to be used, usability testing is an important part of their development.

My study examined how participants ($n = 7$) approached and completed several complicated tasks using Ocean Navigator via a think aloud participant observation method paired with the well-established Likert scale tool, the System Usability Scale. Thereafter, I analyzed this data to determine where Ocean Navigator's usability can be optimized.

1.2 OVERVIEW

In the next chapter, I describe related literature including usability guidelines related to Ocean Navigator, and usability evaluation techniques that will be used in this research. In chapter three, the methodology will be explained, focusing on a description of the participants, a list of materials, a description of the procedure, and an explanation of the data cleaning. Chapter four is where the results will be presented, organized by their relevance to the research questions. In chapter five, the findings will be contextualized using the previous literature discussed in chapter two, limitations will be discussed, and summaries of recommendations for immediate implication and recommendations for future research are presented.

CHAPTER 2

LITERATURE REVIEW

There are three topics that will be covered in this literature review. It begins with general usability guidelines, moves on to mapping application specific usability guidelines, and concludes with usability evaluation techniques related to those used in the research presented in this paper. There are numerous usability guidelines for all kinds of applications, therefore I will include only those that likely pertain to Ocean Navigator.

2.1 GENERAL USABILITY GUIDELINES

The value of usability guidelines is dependent upon who is applying them (Johnson, 2014). They are important but not rigid (Bailey et al., 2006). Applying usability guidelines requires difficult decision-making and prioritization on the part of designers (Johnson, 2014). Sometimes designers will have to decide between two conflicting guidelines, prioritizing the needs of the users for whom they are designing (Bailey et al., 2006). Before thinking about specific usability guidelines there is one parameter that should be maintained: “Don’t make me think” (Krug, 2014, p. 44). That is, users should be able to easily understand how to achieve their goals. They shouldn’t have to think about the user interface.

One aspect that is of particular importance when considering how to develop a user-friendly system is vocabulary. Specifically, vocabulary should fit the task (Johnson, 2014). This kind of focused vocabulary can come up in interviews with and observations of typical users. Hennig et al. (2016) indicate that when presenting information, technical jargon should be

avoided as much as possible. Jargon is defined as “the technical terminology or characteristic idiom of a special activity or group” (Merriam-Webster, n.d.). An example of jargon is the word *stat*. “This is a term, is usually used in a medical context, that means immediately” (Nordquist, 2020), whereas in a mathematical context, *stat* is short for *statistic*.

Accordingly, vocabulary should be familiar to the user to facilitate feature-driven reading as opposed to context-driven reading (Johnson, 2014). Feature-driven reading is a bottom-up process “which stipulates that reading begins with the perception of basic features, such as letters, and ends with comprehension” (Brumberger, 2004, p.16). Context-driven reading is a top-down process “which stipulates that readers map information onto a foundation of prior knowledge as they read” (Brumberger, 2004, p.16). Context-driven reading can occur when the brain encounters something it does not recognize, like jargon, and it is much less likely to become automatic in the way feature-driven reading does, leading to lower comprehension. Panchaud and Hurni (2018) found that there was some confusion around the language used within the interface they were testing and suggest that defining main vocabulary early on and having it accessible through some kind of reference tool usable throughout the application could solve this problem.

Moreover, vocabulary should be consistent (Johnson, 2014). “Never use different terms for the same concept, or the same term for different concepts. Even terms that are ambiguous in the real world should mean only one thing in the system” (Johnson, 2014, p. 293). Such guidelines surrounding vocabulary are important to consider while assessing the usability of Ocean Navigator. While vocabulary, and as such text, is one method of conveying information

efficiently, it may not be sufficient in effectively communicating large amounts of information. With tutorials, for example, users tend not to use them if they are text heavy (Hennig et al., 2016). In such cases, illustrations with concise textual enhancers can be helpful to users (Resch & Zimmer, 2013).

Nevertheless, when text is the only option (e.g. data entry labels, or captions), it is imperative to determine an optimal balance between communicating the necessary information while avoiding the use of superfluous text (Resch & Zimmer, 2013). Sometimes it can be beneficial to use groupings with generic terms which, when clicked, reveal terms below them in a hierarchical fashion, although it is important to limit how many levels are included in order to maintain clarity (Resch & Zimmer, 2013). This kind of visual hierarchy helps people distinguish between what is related to their goals and what is not, helping them to find what they are looking for faster (Johnson, 2014). Specifically, a visual hierarchy:

1. Breaks the information into distinct sections and breaks large sections into subsections.
2. Labels each section and subsection prominently and in such a way as to clearly identify its content.
3. Presents the sections and subsections as a hierarchy, with higher-level sections presented more strongly than lower-level ones. (Johnson, 2014, p.84)

As Ocean Navigator attempts to convey a great deal of information, it will be important to consider how its information is presented. Although a great deal of the information is presented in map form, if some sections of the system are text-heavy, it might be wise for developers to implement the strategies previously discussed.

Sometimes, text really is the best way to present information; for example with data entry labels, or captions. The guidelines surrounding vocabulary discussed earlier in this section are important to keep in mind in this case, as well as other guidelines specifically for data labels. Consistency is important here as well, since inconsistent labels can slow user performance by 10 to 25 percent (Mahajan & Shneiderman, 1997). Inconsistencies include using one label on one page and a synonym on another page when the function is the same, variant capitalization between uses of the same label, using a spelled-out label sometimes and an abbreviation other times, spelling errors, etc. (Mahajan & Shneiderman, 1997). Data entry labels should be clear and concise and consistent in order to assist user comprehension of the data entry (Smith & Mosier, 1986). These should be distinct from, but close to, the data entry field; this can be achieved through visual cues, for example by bolding the label (Galitz, 2007). The data label should either be to the left of and horizontally aligned with the data entry field, or above and left aligned with the data entry field (Galitz, 2007).

There are further guidelines about data entry fields apart from their relationship to their data label. Users should only have to enter data once since having to re-enter data increases the risk of errors (Smith & Mosier, 1986). It can be beneficial to include a search history as it has been shown to increase productivity (Resch & Zimmer, 2013). This way if a user needs to change or delete a data entry they would have access to previously entered data. Format restrictions in data entry fields should be stated explicitly in advance as this leads to fewer errors and trials (Bargas-Avila et al., 2011). An example of the restriction does not improve or impede users (Bargas-Avila et al., 2011).

A common data entry field that can and should have restrictions is a date entry field. How users should enter dates depends on what is important the system: accuracy or speed. When accuracy more important, a drop-down menu is best where a user selects the date, month, and year from three separate drop-down menus (Bargas-Avila et al., 2011). However, when speed is more important, an input field is best where the formatting requirements are to the left or within the input field (Bargas-Avila et al., 2011). When choosing the dropdown option for selecting dates or another kind of menu selection, there are some guidelines that should be applied. One is that unavailable options should be grayed out but remain in view (Li, 2017). Removing them removes spatial consistency which can make things more difficult to learn. Adding a description of why the option is unavailable when it is moused over can be helpful. Another important guideline is to avoid very long dropdowns (Li, 2017). Long menus mean the user cannot see all their options at once, and that steering the mouse can become difficult. An additional guideline to not use dropdown menus when typing would likely be faster (Li, 2017). Another guideline is to keep the menu title visible when the dropdown is open (Li, 2017). This helps users remember what they are doing and stay on task. Additionally, it is important to support keyboard input for navigation within a dropdown menu (Li, 2017). This includes up and down arrows to move through the list and typing a letter which brings you to options in the list that start with that letter.

Dropdown menus are one example of a control in a system. A control is defined as an object in a window or dialog box. They include push buttons, scroll bars, pull-down menus, radio buttons etc. There are three things that should be obvious about a control in a system: what it is for, how it is used, and when it has been used (Stone, 2005). If a control has a label, like many buttons do for example, that label should be consistent across the interface (Mahajan &

Shneiderman, 1997). Synonyms should not be used to label two controls that perform the same function. Providing feedback like status information when a control is used helps support the user's progress toward their goal and helps them understand what is happening (Johnson, 2014).

Another type of feedback is error messages. First and foremost, error messages should show very obviously that an error has occurred; lack of error messages is a huge issue (Nielsen, 2001). They should be readable by a non-expert human and explain precisely what error has occurred (Nielsen, 2001). Error messages placed near where the error occurred work best; error messages placed directly to the right of where the error occurred are rated as the most satisfying and intuitive (Seckler et al., 2012). A simple message that no results have been found is unsatisfactory to most users and if possible, a list of suggested alternatives is preferred (Resch & Zimmer, 2013). "Established wisdom holds that good error messages are polite, precise, and constructive. The Web brings a few new guidelines: Make error messages clearly visible, reduce the work required to fix the problem, and educate users along the way" (Nielsen, 2001). Polite error messages shift the blame from the user to the system. While a system's performance dominates a users' assessment of a that system, apologetic statements make users feel more positive about their interactions (Tzeng, 2004). Apart from error messages, other types of feedback include busy indicators and progress indicators (Johnson, 2014).

When designing an interface that involves pointing and clicking, there are two laws to keep in mind. The first is Fitts' law which says that you can point to a target faster if the target is large and close to your starting point (Fitts, 1954). The second is the Steering law which says that if there are constraints in the pathway from the starting point to the target, you can point to a

target faster if the pathway is wider (Accot & Zhai, 1997). This section covered general usability guidelines that can be applicable to most online systems.

To sum up, several aspects of application usability may inform how well Ocean Navigator works for the general user. There are guidelines surrounding many aspects including vocabulary, large amounts of text, data entry fields, etc. Now I will turn to guidelines that are more specific to mapping applications.

2.2 MAP APPLICATION USABILITY GUIDELINES

The usability of mapping applications has tool-specific guidelines that may apply to web-based mapping applications like Ocean Navigator. Some researchers have investigated the functions and elements within a map application, and user preference regarding both. Hennig et al. (2016) found that participants indicated a preference toward a structure with no nested control elements. They also found that only relevant functions should be displayed in order to reduce information overload, and the most important functions should be accessible without scrolling or clicking a button first. Participants also indicated that including map tools within the map decreases readability and makes the map more confusing. Positioning these tools in a toolbar resolves this issue quite simply. Resch and Zimmer (2013) concluded that pop-up windows should be avoided at all costs, as many users indicate that they close them as soon as they pop up and most of the time before they have even completely loaded. Additionally, advertisements should be kept to a minimum and interactive advertisement should be avoided completely. These general guidelines for mapping applications can increase usability, but more specific functions like zoom have their own specific guidelines.

You, Chen, Liu and Lin investigated what users prefer with regards to the zoom function in a map. Results from this research indicated three main findings: (1) an original center zoom design, one where “the center of the map remains the same as before a zoom” (You et al., p.18), is more effective than a re-center zoom design, one which “uses a user-selected point as its new center to re-frame the zoomed section” (You et al., p.18); (2) an original-center zoom design was better associated with the meaning of the zoom icon, a magnifying glass, than a re-center zoom design; and (3) “a re-center zoom design is preferred to an original-center zoom design in the compatibility of movement directions between the map and user’s mental cognition” (You et al., p.24). Resch and Zimmer (2013), who investigated web-based geo-portals, discovered some user preferences as well. They found that in the map itself, participants indicated a preference to zooming by double clicking instead of choosing from a list of zoom levels. Participants also indicated a preference to a drag-and-drop technique for panning as opposed to navigation buttons on the edges of the map expressing that they found the buttons made the task of navigating difficult and tedious.

These guidelines apply broadly to most mapping applications, but there are some guidelines specifically for applications that utilize bivariate data. Ocean Navigator gives users the opportunity to visualize two variables on the same map. There are some aspects of usability to consider with this. Choosing the right symbols for variables is important, especially with bivariate maps (Olson, 1981). Some symbol pairings lead to a perception of two related datasets, some pairings lead to a perception of two distinct datasets, and some allow the map reader to perceive both depending on the task at hand (Nelson, 2000). An example of a symbol is a line. Different line styles have an effect on how people perceive information presented to them

through a map (Kubíček et al., 2017). Previous research found that people can understand lines with different colours and sizes more quickly than lines with different lightness/darkness and shape (Kubíček et al., 2017). It is easier for people to understand a correlation when two variables are on separate maps, but both ways can get the information across (Olson, 1981). Explanatory notes and a noticeable and clear legend can make things even more comprehensible (Olson, 1981).

To sum up, several aspects of application usability may inform how well Ocean Navigator works for the general user. There are general usability guidelines that will be important to consider. There are also guidelines that are more specific to mapping applications. Before discussing the results, however, it is important to understand how usability can be evaluated.

2.3 USABILITY EVALUATION

Usability is the degree that a typical user can make use of a system to achieve their needs (Stone et al., 2005). Following guidelines can have a tremendously positive impact on usability. Evaluating usability is key in ensuring the guidelines applied have the effect they were intended to. There are many different ways to assess usability including qualitative and quantitative methods. This paper focuses on a qualitative speak-aloud participant observation method, and the well-known quantitative assessment tool: The System Usability Scale (SUS).

Krug (2010, p. 41) defines usability testing practically as “watching people try to use what you’re creating/designing/building (or something you’ve already created/designed/built), with the intention of (a) making it easier for people to use or (b) proving that it is easy to use.” Actually watching users try to navigate a digital environment is different from other tests, such as surveys. Researchers can make in situ observations instead of asking a user for their interpretation of their experience. This can lead to drawing more objective conclusions. His book, *Rocket Surgery Made Easy* (Krug, 2010), describes a usability testing method that involves giving participants a series of tasks, asking them to describe out loud their thoughts while attempting those tasks, and recording audio and screen capture of the entire process. This method is widely used to study the usability of web-based applications (Williams & Foster, 2011; Sonstebly & DeJonghe, 2013; Oakley & Daudert, 2016).

A theoretical model that is of particular utility when examining usability is the mental model. That is, a model in a person’s mind that represents their understanding of how something works; it helps them to predict the results of their actions (Norman, 2013). When a person is presented with a new situation, they need to create a new mental model to help understand that situation (Norman, 2013). When a new situation resembles a situation a person has encountered previously, they might try to approach the situation in a similar way; this is called knowledge transfer (Wikström, 2017). The user takes what they know and uses that to define their expectations and understanding of something unknown. This can lead to the problem of overgeneralizing which is when a previous mental model causes a person to choose the wrong action because it worked on a previous system (Johnson, 2014). Since Ocean Navigator is a new

system to most users, it will be interesting to see what previous mental models they apply in attempts to navigate it.

The SUS is a tool that can be used to evaluate the usability of many different types of applications (Bangor, Kortum, & Miller, 2009) including applications like Ocean Navigator. It is a freely available ten-item, Likert-scale survey that helps quantify the inherently qualitative study of usability (Brooke, 1996). It yields a score between zero and 100 (Brooke, 1996). A 100 point scale can be intuitive since it allows for relative comparisons (Bangor, Kortum, & Miller, 2009), but the SUS does not translate well to a percentage score (Brooke, 2013). In a 2009 study (Bangor, Kortum, & Miller), an adjective rating scale was added at the end of the SUS that allowed participants to rate the usability of an application from “worst imaginable” to “best imaginable” (Figure 1).

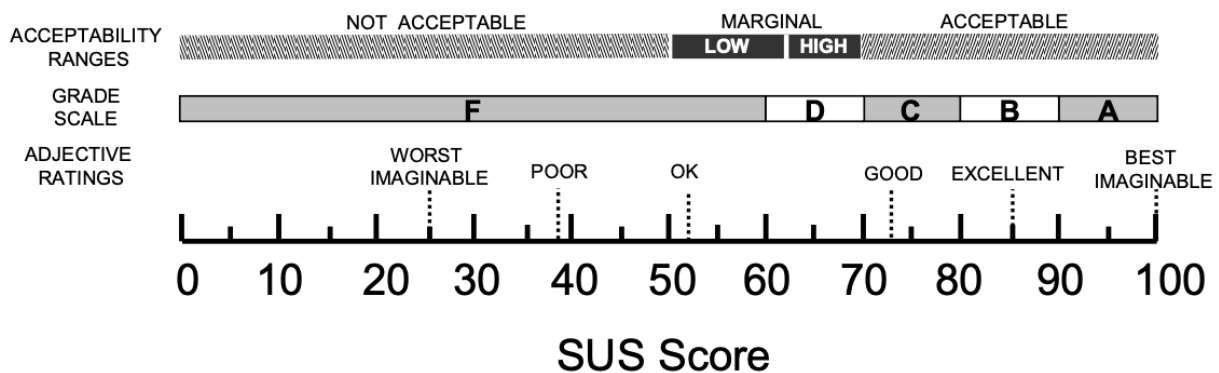


Figure 1 A comparison of the adjective ratings, acceptability scores, and school grading scales, in relation to the average SUS score. Source: (Bangor, Kortum, & Miller, 2009, p. 121)

These adjective ratings can help people understand the numerical score the SUS produces. The SUS is considered to be reliable and valid, but it is not considered to be a diagnostic tool (Sauro, 2011). It was created to be a quick and easy way to determine people’s satisfaction with an

application, and over the years it has proved to be able to do just that (Brooke, 2013). Therefore, I asked participants to complete tasks on Ocean Navigator while thinking aloud and to fill out the SUS so I could answer the following questions:

1. Are participants able to use Ocean Navigator to meet typical information needs of the expected users? How does the completion of tasks by novice users compare to that of an expert user?
2. Do people perceive Ocean Navigator as usable? How?
3. Do users translate their expectations from consumer tools to niche tools like Ocean Navigator?

CHAPTER 3

METHODOLOGY

A mixed-method research design was employed in order to provide a holistic representation of the usability of Ocean Navigator. Both qualitative and quantitative data were collected. Quantitatively, a descriptive research design was employed using a survey. In this study, descriptive research means that this data helped describe the phenomenon being studied. This portion of the data helped to answer the “what” aspect rather than the “why”. This data was collected through the Systems Usability Scale (Brooke, 1996) (SUS) (see Appendix E), a widely used series of Likert scale questions that attempt to quantify the usability of a software system. A study by Bangor, Kortum, and Miller (2008) calls the SUS “a highly robust and versatile tool for usability professionals” (p. 547) that “can be used to positively supplement a usability testing and evaluation program” (p. 590). SUS ratings were completed by participants after finishing the tasks described in the following paragraph.

Qualitatively, a phenomenological research design employed a speak-aloud method of interview. This portion of the data helped to answer the “why” aspect for the “what” that was identified in the quantitative. This data was collected through a participant observation method explained thoroughly in the book *Rocket Surgery Made Easy* (Krug, 2010). In this method the user is given a series of tasks (Appendix D) and asked to describe out loud their thought process as they complete those tasks. Screen capture and audio recording were collected. After the interviews were transcribed, they were analyzed using content analysis. Content analysis involves familiarizing oneself with the data you are analyzing, generating codes from ideas that emerged through that familiarization, searching for themes among the codes, reviewing themes

to help the data tell a story, defining and naming themes, and performing final analysis and producing a report (Braun & Clarke, 2006). Atlas.ti software was used to assist in the storage and creation of thematic codes. Screen capture was used to determine if tasks were completed correctly. This method is widely used to study the usability of web-based applications (Williams & Foster, 2011; Sonsteby & DeJonghe, 2013; Oakley & Daudert, 2016).

The qualitative data was used to explain and give further meaning to the results of the quantitative data. The SUS will produce a single score for each participant. The qualitative data will give us a better idea of that that score really means both for participants individually, and for the group as a whole once the scores have been averaged. Mixed methods can be beneficial when researching new areas because of this ability of quantitative data to describe a phenomenon, and then qualitative data to help give meaning to that description. The qualitative data was collected first to ensure that participants had experience with the system when they were rating it, and didn't base the rating on first impressions

Cognitive fatigue is a state that can occur after extended involvement with cognitively challenging tasks (Meijman, 1991). Acute cognitive fatigue can lead to decreased motivation (Ackerman & American Psychological Association, 2011). Despite fatigue impacting motivation, research has shown that it does not impact SUS ratings by participants (Robertson & Kortum, 2017). This decline in motivation could however be more apparent in the qualitative data.

3.1 PARTICIPANTS

Ocean researchers from across Canada gathered together for Marine Environmental Observation Prediction and Response (MEOPAR) Network's Annual Training Meeting and Annual Scientific Meeting in Victoria, BC, Canada, in June of 2019. Volunteers recruited from the conference attendees who partook in the speak-aloud participant observation method, as well as the SUS method, were considered 'first stream participants'. All students and postdoctoral fellows attending the conference were invited, but response was lower than anticipated. These participants were chosen because they were new to Ocean Navigator but had familiarity with the language and concepts that would be presented. There were eight participants in total in the first stream. They included different types of users such as graduate students and post docs. As an incentive, ten dollars was offered as either cash or a donation to an ocean-related charity. Second stream participants were recruited vigorously through email. Students from multiple schools with oceans programs, members of the Marine Environmental Observation Prediction and Response Network, members of Fisheries and Oceans Canada, etc. were contacted for recruitment purposes. They participated remotely through Opinio, a web-based survey software. There were two participants in total in the second stream. They partook in solely the SUS method and no screen or audio recording was collected. This is partially because we were expecting at least 30 participants in this section, and partially because asking people to download and use screen recording software is unreasonable considering the amount of support they would require. Since this is one of the first studies of this type on a tool like Ocean Navigator, we kept our study design to a small sample, exploratory assessment of the system to see what areas deserve more attention and to get an initial sense of user reaction to this tool.

3.2 MATERIALS

Materials necessary for testing with first stream participants were a script which was adapted from *Rocket Surgery Made Easy* (Krug, 2010) (see Appendix A), a consent form (see Appendix B), recruitment emails (see Appendix C), task descriptions which were provided by Fisheries and Oceans Canada, then reviewed, simplified, and clarified by the research team (see Appendix D), SUS Likert scales (see Appendix E), incentive (\$10.00), and a computer with screen capturing and sound recording software. Materials necessary for testing with second stream participants were an Opinio survey administered online which included a consent page (see Appendix B) and the SUS Likert scales (see Appendix E). Recruitment emails (see Appendix C), and task descriptions (see Appendix D) were also used.

3.3 PROCEDURE

First stream participation took place in a room at the Delta Hotels by Marriott Victoria Ocean Pointe Resort in Victoria BC. I read a script (see Appendix A) instructing participants through the process. The process included briefing participants to the purpose of the study, signing a consent form (see Appendix B), acclimation to the site, completion of tasks (see Appendix D), completion of SUS (see Appendix E), and debriefing. Screen recordings and sound recordings were captured, using Panopto, showing what the participant was doing and saying during the tasks. The laptop used was a laptop from our research lab, with Windows 7 installed and running the Google Chrome web browser. An external mouse was attached to avoid concerns of unfamiliarity with a laptop trackpads. Second stream participation took place

remotely through an Opino survey distributed by email. These participants were asked to familiarize themselves with Ocean Navigator with the suggestion of completing the tasks (Appendix D) and completed the SUS rating scales (Appendix E).

The audio recordings were transcribed and then organized by task. I was striving for verbatim transcription because pauses and hesitations are significant, and quotations appearing in the thesis are similarly verbatim. Following this transcription and superficial organization, I followed Braun & Clarke's (2006) method and preliminary categories were identified from themes that emerged while transcribing to help in coding the data. The categories included comparison, suggestion, emotions, and task status. Subcategories were then identified as appropriate. For example, the category emotions included such subcategories as frustration, satisfaction, confusion, etc. Following the development of a list of categories, or codes (See Appendix F), the transcriptions were divided into phrases, sentences, and sections of text which were coded individually. The initial codes were then applied to the data. While applying the initial codes, new codes were discovered (see Appendix G). After the first round of coding the data, codes were added and renamed as necessary (see Appendix H) producing a second list of codes and subcodes (see Appendix I) which were applied to the data in round two of coding. Following this round, definitions (see Appendix J) were created that were as specific and concrete as possible. Again, as suggested by Braun and Clarke (2006) the codes were applied a final time in round three of coding using these revised definitions.

The coded data was then analysed with the research questions in mind. For the first research question, (Are participants able to use Ocean Navigator to meet typical information

needs of the expected users? How does the completion of tasks by novice users compare to that of an expert user?) I noted which tasks were incomplete and looked at codes that were applied in those cases. They also noted which tasks were complete, but participants struggled with, and looked at codes that were applied in those cases. In addition to the audio recording/screen capture of the participants, an audio recording/screen capture of one of the developers of Ocean Navigator was collected as a baseline. The participants' performance was compared to that of the expert user.

The second research question (Do people perceive Ocean Navigator as usable? How?) was analyzed by looking at the SUS results. The SUS has an already established method of analysis that involves coming up with a score that represents the usability of the tool being examined. This method of analysis yields a single number representing a composite measure of the overall usability of the

...system being studied. Note that scores for individual items are not meaningful on their own. To calculate the SUS score, first sum the score contributions from each item. Each item's score contribution will range from 0 to 4. For items 1,3,5,7, and 9 the score contribution is the scale position minus 1. For items 2,4,6,8 and 10, the contribution is 5 minus the scale position. Multiply the sum of the scores by 2.5 to obtain the overall value of SU. SUS scores have a range of 0-100 (Brooke, 1996, p. 5)

The SUS score allowed us to utilize descriptive statistics, such as the mean SUS score. Qualitative data helped give nuance to the quantitative data collected in the SUS by looking at codes that were applied in relation to participant's perception of the usability of Ocean Navigator.

Research question three (Do users translate their expectations from consumer tools to niche tools like Ocean Navigator?) was analysed by looking at when the code “COMPARISON” (Appendix J) was applied, which translates to when participants compared Ocean Navigator to another system. When it was applied, insight was gained by looking at what applications participants were referring to, when they referred to those applications, and whether they were saying something positive or negative about Ocean Navigator through the comparison.

3.4 DATA CLEANING

Eight participants completed the usability testing. However, one participant was excluded from analysis. Participant three’s level of comprehension would make them an extreme outlier and thus they were removed from the analysis. They seemed to struggle with both English language comprehension and the kind of information presented on Ocean Navigator. They required substantially more guidance in completing tasks which would make it difficult to tell if the participant was having trouble with the usability of Ocean Navigator, or if the trouble was with language comprehension or the kind of information presented on Ocean Navigator. Additionally, it would be difficult to separate their understanding of the task from the guidance I provided so their responses might just reflect the guidance provided as opposed to the usability of Ocean Navigator. It was decided not to analyze the data gathered from the two second stream participants because response was so much lower than anticipated.

Participant responses to all questions and tasks were used in determining whether participants had completed a task except for responses to Task 3a. Task 3a was omitted from this analysis because the expert user ignored the date requirement making it impossible to compare to the participants' performance. However, some qualitative analysis was still possible because participants still attempted or completed the task and gave valuable verbal feedback while doing so.

4.1 DOES OCEAN NAVIGATOR MEET THE TYPICAL NEEDS OF THE USER?

Participants are able to use Ocean Navigator to meet the typical information needs of the expected user. If participants were able to complete tasks, it was inferred that they were able to use Ocean Navigator. Five of the tasks, Tasks 1a, 1d, 2a, 2b, and 2c, were completed by 100% of the participants. However, some aspects of the program do not provide a good user experience. Three of the tasks, Tasks 1b, 1c, and 3b, were completed by less than 100% of the participants. Some tasks were completed by 100% of participants, but some participants received assistance to complete them in order to keep the session progressing.

While all the participants completed Task 1d, there was some difficulty. Specifically, participants struggled with drawing a line and jargon. When asked to use the “draw a line” function and retrieve data from that line, participants struggled with ending their lines. Participant 7 said, *“I need to kill that how do I do that escape kill that I need to end this line how do I do it escape I don't know maybe double-click”*. Participant 2 attempted to end the line by clicking once and then said, *“oh, double click?”* and successfully ended their line. Participant 5 similarly attempted to end the line by clicking once and then said, *“if I double-click? Okay”*. and successfully ended their line. A further difficulty within Task 1d was that participants did not all know what a Hovmöller diagram was. A Hovmöller diagram is a mixture of a graph and a map that shows movement in static pictures (Hovmöller, 1949). The diagram shows the average of all the values in a single row of latitude or column of longitude and display those values on one axis, with time represented on the other axis. Dialogue with participant 6 was as follows:

Participant: “the intended audience for this is who”?

Researcher: “ocean researchers [who] wanna look at data I guess”

Participant: “so people who know what a Hovmöller diagram is”

Researcher: “likely yeah probably you wouldn’t be trying to make one if you didn’t know what it was”

Participant: “just thinking like pretty it’s jargon if we’re not”.

Participant 8 expressed their confusion as well saying, “*but I don’t know what a Hovmöller diagram is*”.

While Task 2 was completed by all participants, in many cases participants struggled with 2a and help was provided in order to keep the test progressing. Task 2 involved comparing data observed in situ to a predicted model using Class 4 metrics. Class 4 is a niche term used only by ocean modelers, and not the rest of the ocean research community. Dialogue with participant 7 during task 2a was as follows:

Researcher: “and I want you to look at this top bar”

Participant 7: “okay, oh Class 4 I’m supposed to pick right because you’ve told me that I’m making a Class 4 metric so I guess to ‘visualize’¹ my feelings”

Researcher: “mm-hmm”

Participant 7: “I guess because I’m not like familiar with the term Class 4”.

Participant 6 also brought up the issue of not knowing what Class 4 meant saying, “*I feel like there’s an assumption that you know what Class 4 is*”. during their closing remarks. Participant 7

¹ The direct quote uses the word visualize but given the context it is likely the participant meant verbalize

said, “I'm surprised that I, given my field of work, I'm not familiar with it but anyway okay”.

Question 2a took participant 7 6.1 times as long as the expert (Table 1).

Table 1: Time (ratio) participants took to complete tasks relative to an expert user.

Tasks	P01	P02	P04	P05	P06	P07	P08	Expert	Average
1a	2.3	5.2	2.0	4.5	5.4	3.2	3.2	1	3.7
1b	2.6	2.7	2.1	3.2	2.0	4.7	2.1	1	2.8
1c	4.7	2.0	1.5	3.6	1.8	1.9	1.3	1	2.4
1d	4.7	6.1	4.4	4.8	4.4	4.5	5.4	1	4.9
2a	3.4	7.4	3.9	4.2	2.0	6.1	4.6	1	4.5
2b	0.4	2.2	0.4	0.4	0.3	1.1	0.6	1	0.8
2c	1.0	1.5	1.2	1.0	0.8	1.5	0.8	1	1.1
3a	3.1	2.4	0.9	1.0	1.5	2.4	1.9	1	1.9
3b	--	12.2	1.4	1.0	2.2	0.8	2.6	1	3.4
Average	2.8	4.6	2.0	2.6	2.3	2.9	2.5	1	2.8

However, participant 7 completed tasks 2b and 2c with times much more comparable to the expert, 1.1 and 1.5 times as long respectively. Participant 7 indicated that they felt they had completed 2c saying “okay but I think I have successfully visually compared the temperature profiles as requested all right”. This difference in completion times for Task 2a and Tasks 2b and 2c was apparent in all participants. It took participants an average of 4.5 times as long as the expert to complete Task 2a. However, it took 0.8 and 1.1 times as long to complete Tasks 2b and 2c respectively (Table 1). Four participants (P4, P5, P7, and P8) needed help in order to find the Class 4 button in task 2a, but no participants needed help for tasks 2b and 2c. Tasks 2a and 2b/2c were similar so once participants figured out the Class 4 function in 2a, they were able to complete 2b and 2c much more quickly since they did not need the time to figure out the Class 4 function again.

One of the tasks that was completed by less than 100% of the participants was Task 1b. When asked to use a virtual mooring to access monthly average water temperature at the bottom of the Atlantic Ocean, participants struggled for a few reasons. Struggles faced by participants included difficulty finding the bottom option in the list of depths, difficulty entering dates, lack of error messages, and lack of memory in the system. There was a list of numerical depths that went deeper and deeper the further down the list you scrolled with the exception of the bottom option being at the top of the list. When struggling to find the bottom option, participant 7 said:

...so now I feel conflicted cuz I want the depth here and now I'm worried that the depth listings in this menu are not necessarily specific to my location right so how do I know what the bottom is besides trial and error?

After selecting a depth they estimated to be the bottom, participant 1 said, *"I'm still not a hundred percent sure if that's the bottom; I'm gonna try to pick one the depth of greater and see if it still generates a plot"*. Participants also struggled with entering dates. Many attempted to type the date in instead of selecting it from the calendar, which took a long time to click through. Participant 7 verbalized this frustration saying, *"it would be nice if I could type this into so I don't have to click it all these many times"*. Participants also struggled with the lack of error messages. Participant 8 asked, *"Does it sort of freeze up like this sometimes or is there something I'm doing wrong"*? Participant 7 said, *"yeah I continue to be confused about this whether is this data available or is my internet slow"*. Participant 2 was also not sure if the internet was to blame saying *"let's blame the wifi of the hotel"* while looking at a loading screen that never loaded. Another aspect participants expressed dissatisfaction with was that the system does not remember what you have previously typed. When entering coordinates, if you get one wrong and go back to change it, you have to enter all the coordinates again. Participant 7 voiced

their frustration saying, *“I want to enter the coordinates; I got to do this again, I feel frustrated... three five, maybe it could remember where I’ve searched before”*.

Another task that was completed by less than 100% of the participants was Task 1c. When asked to create a map of a given area that showed two variables, most of the participants struggled. Only 3 of the 7 participants completed this task successfully (Table 1), and those 3 took approximately twice as long as the expert user. Up to that point all of the tasks they had seen visualized a single variable, but task 1c required using color to visualize the temperature, and arrows to visualize ocean current (or velocity). Participant 4 thought they were supposed to add a velocity contour and said, *“Average ocean current or water velocity would be an additional contour but which one, X or Y”*? Participant 5 created two maps, one showing temperature and one showing ocean current but did not know how to create one map showing both. They said, *“Okay so here I have temperature and water velocity I don't know, like, how can I combine both of them”*? Even participants who completed this task successfully in the end struggled to add a second variable. Participant 8 said, *“oh I need average surface temperature and average ocean current... so I'm not sure if temperature by default (mumbles) September 2018... I'm just not sure how to add second”*.

Users prefer short term gain over long term benefit when learning to use a new tool. Task 3b was completed by less than 100% of participants, possibly for this reason. Task 3b asked participants to download an R script to generate a plot in R that they had just generated in Ocean Navigator. Three participants were unsuccessful. The participants who completed this task did so with times comparable to the expert (Table 1) which could suggest that it was not the difficulty

of the task or the usability of Ocean Navigator that was the problem, but rather that participants wanted to be done and it was the last question. Participant 2 verbalized this feeling saying, “*And I just, no, I’m too lazy now to (laughter)*” and while the other two participants who were unsuccessful did not verbalize this feeling it was apparent in their non-verbal vocalizations and body language. Participant 1 even seemed to not notice the question saying, “*I think I’m just gonna call it a day*” after feeling unsuccessful with task 3a. The four other participants indicated satisfaction with task 3b. Participant 4 said, “*okay so I’m just saving that, and I think that is all*” and participant 6 said, “*okay cool so we’ve got an R file*”.

Although task 3a was removed from the quantitative analysis, there were findings in the qualitative data. When asked to upload a CSV file to access data from the area defined in the CSV, participants struggled with clicking the right area to access data after uploading the CSV. Dialogue with participant 1 was as follows:

Researcher: “yeah you were clicking in the center which maybe it would be helpful to make that a clickable area as well because I find you have to click right on the line and that might not be the best way to do it”

Participant 1: “definitely did try clicking on the center and ‘left²-clicking on the center”.

Participant 1 also said “*I just didn’t know that I need to do that to launch the screen*” when talking about clicking the area. Participant 4 suggested that it might be easier if the area were more obvious. Dialogue with them was as follows:

Participant 4: “I wonder if there can be like a bit more of a kind of like a highlighting or something on it”

² The direct quote uses the word left but, judging by what the participant did, it is likely they meant right

Researcher: “yeah”

Participant 4: “just because it pops up for some things but then doesn’t”

when referring to a screen that displays data.

4.2 USERS’ PERCEPTION OF OCEAN NAVIGATOR AS USABLE

The SUS yields a single number representing a composite measure of the overall usability of the system being studied. With an average SUS score reported by participants was 72.69 with a standard deviation of 8.49 (Table 2), users perceive Ocean Navigator as usable, but not extremely so.

Table 2: Participant’s SUS score and average completion time relative to an expert user.

Participant	SUS score	average completion time relative to expert
P1	75	2.8
P2	85	4.6
P4	67.5	2.0
P5	60	2.6
P6	81.25	2.3
P7	70	2.9
P8	70	2.5
average	72.68	2.8

This score means that Ocean Navigator is between “OK” and “GOOD” according to Bangor, Kortum, and Miller (2009), a group who researched interpreting SUS scores (see Figure 1 in section 2.3). There is a positive correlation of 0.58 between SUS scores and completion time relative to expert. While the sample size is too small to achieve statistical significance ($p=.128$), this means that the longer it took participants to complete the tasks, the higher they rated the

system on the SUS. In addition to this quantitative view on the usability, dialogue throughout testing garnered some insight into usability as well.

Qualitative data about the dataset information, an aspect some participants felt did not provide good usability, adds nuance to the SUS score achieved by Ocean Navigator. Some participants did not notice that dataset information was available. Participant 1 said, *“maybe if there was a page where you could have a little bit of a description of the datasets, but I could always just Google that in a sidebar which is normally what I would do”*. Some participants clicked it and then closed the window quickly after noticing the length of it, including participant 7 who said, *“I guess I could read this thing but it's long and I don't want to read it all”*. Participant 7 seemed highly concerned with the source of data and had a lot of things to say about it including, *“but is the monthly date is that really a different dataset or just a different averaging of the same dataset”*, *“so is this just a model or does it is it an assimilated model that has data with it I think because it doesn't say anything about data it's just a model”*, *“I know nothing about those observations like where did they come from what is their source so I would be curious about that”*, and

I want to know what climatology is that is that the world ocean atlas or like... and climatology over what time period yeah so I guess it would be nice to be able to like click on a information and get more information both about the observations and maybe it's impossible like it's unfeasible but ya know?

The dataset information was invisible to some participants and too much information for others.

Another aspect of usability that may have affected the SUS score was selecting start and end dates. Many participants struggled with the fact that the system will not let you select start and end dates that are incompatible with each other. Dialogue with participant 7 was as follows:

Researcher: “you didn’t save the first one so it’s not letting you go back further than that”

Participant 7: “oh so I didn’t actually select April 2016”?

Researcher: “yeah”

Participant 7: “okay so I wasn’t aware of that yeah and so now I’ve got to click again, is there a better way to do this”?

Researcher: “uhh there could be”

Participant 7: “April 26”

Researcher: “you didn’t select it”

Participant: “okay so now I feel frustrated”.

Participant 2 noticed this issue as well and pointed that out saying, “*the end time is after start times, it could be something*”. Participants found the process of selecting dates tedious even when they did know how to select them properly, making sure start and end dates were compatible. Participant 8 said “*there’s probably an easier way to change this year*” while clicking through the years one by one. Participant 6 expressed frustration with the system for going through every date they had clicked through, loading each one, instead of going to their final selection. Dialogue with them was as follows:

Participant 6: “first of all, not efficient, B: this is also going to be extremely inefficient probably very very shortly because”

Researcher: “it’s gonna jump through all the dates”

Participant6: “(laughter) it’s gonna try to jump through all the dates”.

They had wanted to change from two dates that were very far apart and clicked through many dates before arriving at their desired date. When they arrived at their desired date, Ocean Navigator loaded data for every date they had clicked through and did so at a slow pace.

How long tasks took to complete is another usability aspect that participants were not satisfied with. A few participants commented that they felt the tasks were taking a long time which decreased their perception of the usability of Ocean Navigator. Participant 2 blamed the system saying, *“and since I have to share my thoughts it’s it’s a little bit long but I understand it’s a lot of data so maybe it’s...”*. Alternatively, three participants blamed themselves. Participant 8 said *“I thought it was easy to use again bearing in mind it took me a long time to do that but I’m not used to any similar system”* while answering the question “I thought the system was easy to use” from the SUS. Participant 4 and 1 expressed that they thought they were taking too long on the tasks saying, *“has it usually been taking this long”?* and *“I spent a little long on that”* respectively.

Adding a help page can increase usability, but a blank help page can leave participants feeling confused. Participants were upset that the help page on Ocean Navigator was blank. Participant 5 said, *“I don’t have any help? Come on”* while trying to access the help page and brought it up again in their closing remarks saying, *“the help didn’t work”*. Participant 2 said, *“maybe I could ask help? No (laughter) I won’t”* while attempting to access the help page and participant 1 had a similar experience saying, *“there’s a help button let’s click that. Compare dataset help? But it doesn’t say anything, okay, hmm”*.

When participants were completely stuck, I provided help in some cases which could have affected the participants' perception of the usability of Ocean Navigator. Six of the seven participants had to refresh the page at some point during the session. This solution was recommended by me so it's possible that participants would have been unsuccessful with some tasks had they not received that help. I provided assistance to one or more participants in tasks 1a, 1b, 1c, 1d, 2a, and 3a. Help provided was often suggesting the right way to do something to progress the task when the participant was not successful (see Appendix J for a full definition of the code "HELP: researcher provided").

4.3 SOME USERS TRANSLATE THEIR EXPECTATIONS FROM CONSUMER TOOLS TO NICHE TOOLS LIKE OCEAN NAVIGATOR

Two participants referred to other systems while completing tasks on Ocean Navigator and together they referred a number of different application, websites, and/or tools including Adobe³, MATLAB, Python, and worldview.earthdata.nasa.gov. The references tended to highlight the positive aspects of the comparator program.

Participant 6 referred to Adobe when trying to explain his feelings about menu location on Ocean Navigator. Dialogue with them was as follows:

Participant 6: "but I feel like Adobe has trained me to look for menus a majority of menu systems like on one side and I know that it can move my Adobe around but I'm old-school Adobe before I could move it around so I've got them all on the same side always and that's

³ Adobe was mentioned generally, the specific Adobe product was not mentioned

just like a little like expanded collapsy thingy on that side and then like just the select few main tools are on one side but all my menu options that go like extra stuff that you can add”

Researcher: “so like if this collapsed and this collapsed and this collapsed and they were all on one side you think that would be easier”?

Participant 6: “for me”

Ocean Navigator has options along the left side of the screen and going across the top of the screen. Participant 6 would prefer all options to be in the same area instead, similar to Adobe-style applications, and to use collapsible options so there would be room for everything.

Other participants compared Ocean Navigator to MATLAB for how it handles error messages, creating diagrams, and exporting data. Participant 6 referred to MATLAB when discussing what they perceived to be a lack of error messages on Ocean Navigator. They said, *“least favorite features about old MATLAB inaccurate error messages, current MATLAB has much better error messages it's much less frustrating and hair pulling”* after expressing a desire for accurate error messages. Participant 7 also referred to MATLAB but in another context. In Task 1d, when creating their Hovmöller diagram, they said, *“if I was making this in MATLAB myself, I would then draw a contour where the ice concentration is or the ice concentration fraction is 0.15”*. They successfully completed the task but wanted to go further and were unsatisfied with their inability to do so. There is a way to get all of the raw data from Ocean Navigator, but this participant did not realize this despite their strong desire to do so. Participant 7 referred to both MATLAB and Python when discussing their desire to export the data. They said, *“I feel like I don't know I guess because I'm a quantitative person and I work with things*

like MATLAB Python or whatever I want to be able to export that data as like numbers". When asked to explain why participant 7 indicated that they would like to be able to do more with the numbers saying, *"...so that I can plot them myself, I can do a calculation or whatever"*.

Finally, Participant 6 referred to a NASA's Worldview website (worldview.earthdata.nasa.gov/) when talking about selecting dates. They said, *"date selecting is a lot easier"*. This website has arrows to click through dates, but you can also click and type in your date (Figure 2).



Figure 2 Screenshot of the date selection widget from worldview.earthdata.nasa.gov

Participant 6 also made a more general comparison saying, *"yeah I mean that's you can do it pretty easily it's not like it's that hard but I feel like these days we're so used to just like well of course it knows what I want to do I want the next thing right there I don't want to have to select"*. They brought this up while struggling to get a profile to pop up. Participants tended to make comparisons when they found other applications were doing something better than Ocean Navigator.

Many things need to be taken into consideration to put the results in the previous section into the context of the research questions. Some findings were to be expected and make sense when considered with the previous literature in mind. Other findings contradict the previous literature. There are limitations to this methodology including my bias, unexpected events that occurred during the process of the speak-aloud participant observation, the limited sample of participants, etc. These findings add to the body of knowledge surrounding usability, but also bring up a lot of questions for future research.

5.1 FINDINGS

5.1.1 Does Ocean Navigator Meet the Typical Needs of the User?

The typical needs of the Ocean Navigator user would be to investigate and visualize ocean data. Users can obtain an overview of oceanographic data and launch further evaluations of ocean forecasting and comparison.

One of the ways to answer this question is to look for instances when participants did not know or understand something. One point of struggle was vocabulary. There were two tasks that involved terms participants did not know. These terms were “Hovmöller diagram” and “Class 4 metrics”. Another term, “stick plot”, was identified as jargon by Participant 8 while they were trying to figure out how to display virtual mooring data. They said,

[S]o, it looks like it should just be point data so I should be in the right place... This shouldn't matter but I'll just see what it gives me for options... I'm not sure what a stick plot is, I don't know so maybe bottom... I'll just go to the end of the month.

Since the task did not require a stick plot, the participant only mentioned it in passing and then moved on. However, had a task involved a stick plot, this suggests that there could have been issues. The tasks that did involve jargon though, did cause difficulty for some participants. For example it took participants an average of 4.5 times as long as the expert user to complete Task 2a, the task that first introduced the term "Class 4 metrics", but subsequent tasks that involved the term "Class 4 metrics" were completed with times almost identical to the expert user. This is consistent with the literature and there are some simple solutions that could mitigate this problem. One solution would be to use focused vocabulary that is more familiar to users which can be found by conducting interviews or observations with typical users (Johnson, 2014). This option would require significant effort and is a recommendation for future research. Panchaud and Hurni (2018) suggest defining main vocabulary early on and having it accessible through some kind of reference tool usable throughout. This solution would likely be easier to implement and is my recommendation for more immediate action.

Another usability problem area for participants was the dropdown menu for depth selection. Things to consider when creating dropdown menus include, graying out unavailable options, avoiding excessively long lists, keeping the menu title visible, supporting keyboard navigation, and avoiding dropdown menus completely if typing would be faster (Li, 2017). These guidelines for dropdown menus could also be useful in solving the problem that occurred with the depth selection dropdown menu. The depth selection menu was long and did not gray

out unavailable options. However, I propose that this can be addressed by following the guideline that says not to use a dropdown menu if typing would be faster. Instead, use a text input field for numerical data, and since Ocean Navigator has a “bottom” option, which would be less user friendly for typing, a radio button for that option would be an ideal solution.

Another point of struggle that leads to lower usability was drawing a line. When participants were asked to draw a line, a lot of them struggled at first to end their lines. They did figure it out quickly but how a control is used should be obvious (Stone, 2005). There are some suggestions that would make it more obvious upfront. Providing feedback like status information, or progress indicators when a control is used helps support the user’s progress toward their goal and helps them understand what is happening (Johnson, 2014).

Creating maps displaying bivariate data can be difficult and for Ocean Navigator, this was a usability problem area. In Task 1c participants were able to add the temperature variable with little to no trouble but adding the additional variable of velocity was not as intuitive. Choosing the right symbols for variables is important (Olson, 1981). Ocean Navigator used colour for temperature which users understood well. I recommend using arrows that are distinguished by colour and/or size to represent velocity as suggested in previous research (Kubíček et al., 2017). I also recommend including explanatory notes as well as a legend as suggested in previous research (Olson, 1981). Another concern was that the word “velocity” was used both for arrows and contours on Ocean Navigator. The same term should never be used for different concepts (Johnson, 2014). I recommend changing or deleting one instance.

Another usability problem area was clicking a target. Participants struggled with clicking the area after uploading it from a CSV. The clickable area was quite small so this makes sense considering Fitts' law which says that you can point to a target faster if the target is large and close to your starting point (Fitts, 1954). I recommend making the clickable area large for all clickable targets.

Lack of memory was a point of struggle leading to lower usability. Participants were dissatisfied with their experience when entering coordinates, specifically with the issue of Ocean Navigator not remembering coordinates you had already typed. When a user needs to change one coordinate of an area, they are required to re-enter all of the coordinates. This dissatisfaction is in line with the literature which indicates that including a search history can be beneficial (Resch & Zimmer, 2013), and that users should only have to enter data once (Smith & Mosier, 1986). Although the issue did not present itself during this study, another strategy to keep in mind is that format restrictions should be stated explicitly in advance as this leads to fewer errors and trials (Bargas-Avila et al., 2011).

Motivation was another issue. Participants in these sessions had no real stake in the results of the research. This can be a considerable issue with this type of methodology. While an incentive (\$10 in either cash or a donation to an ocean related charity) was provided, lack of motivation is definitely a possibility. However, since we wanted to see things like when lack of motivation kicks in or when a task was too difficult, a bigger incentive was thought unwise. Some participants mentioned that Ocean Navigator could have been or could potentially be useful to them. Participant 8 indicated this saying, *“yeah we definitely coulda used this in our*

projects, it's pretty cool". They were also the only participant to successfully complete every single task, including 3b which was only completed by four participants in total (see Table 1, section 5.1). However, others felt like this tool was unrelated to their research or that they would not be typical users. This lack of motivation could really be seen in Task 3b, the last task of the session. One participant said they were too lazy to finish the task, and another did not even notice Task 3b. In these situations, it felt obvious that the participants were not motivated to complete the session.

5.1.2 Users' Perception of Ocean Navigator as Usable

Measures of whether users perceived Ocean Navigator as usable could be seen in their SUS ratings, as well as some of the dialogue that transpired.

There is a correlation of $r=0.584$ between SUS scores and completion time relative to expert. While this result is not statistically significant, this may suggest that the longer it took participants to complete the tasks, the higher they rated the system on the SUS. It would make sense for participants to find the system less usable if it took them longer, but this is the opposite of what was found. This may suggest that users with more experience working with data systems were both able to finish faster, and had higher expectations thus blaming the system and giving lower scores. This would fit with the fact that two of the three users who finished fastest rated the system lowest (see Table 2). Further study is recommended since there is not sufficient evidence to claim that this is true or that it is generalizable to a larger group of users.

There were two main problems identified through this research pertaining to the dataset information available in Ocean Navigator. One, participants did not notice it existed; and two, when they did notice it, they felt it was much too long, and thus did not use it. The problem of length is consistent with the literature which states that users tend not to use tutorials if they are too text heavy (Hennig et al., 2016). My recommendation for this problem is to implement a visual hierarchy since these can help people distinguish between what is related to their goals and what is not, thus helping them to find what they are looking for more quickly (Johnson, 2014). Resch and Zimmer (2013) suggest using groupings with generic terms which, when clicked, reveal terms below them in a hierarchical fashion. Ocean Navigator could implement this strategy in their dataset information by having the dataset name as the highest level in the hierarchy, and the only thing that appears initially in the list, which when clicked, reveals the information about that dataset.

Participants struggled with selecting dates in Ocean Navigator. The method most participants used was to click through using the arrows on the calendar. This did not work well with the system because Ocean Navigator would load new data for every new date even if the participant was just clicking through them to get to the date they actually wanted. The literature recommends two date selection methods, neither of which involves clicking through multiple dates. Since accuracy is important for retrieving the correct data, as opposed to speed, Ocean Navigator should use three drop down menus where a user selects the date in the first dropdown, the month in the second, and the year in the third (Bargas-Avila et al., 2011).

Since the help page was blank, I cannot be certain about what DFO wishes for it to accomplish. However, I assume it was supposed to house some type of tutorial to assist users in operating Ocean Navigator. Assuming that is the case, my recommendation is not to solely use text, but to use illustrations, or screen captures, with concise textual enhancers as suggested in Resch and Zimmer's (2013) research.

When participants were trying to access data that wouldn't load for one reason or another, they were dissatisfied and confused. Participant 8 asked, "*Does it sort of freeze up like this sometimes or is there something I'm doing wrong*"? Participants did not know why the page was not loading. The literature suggests that not including error messages is a grave omission (Nielsen, 2001) so my first recommendation is to use error messages. Some important guidelines when creating these error messages are that they should be readable by a non-expert human (Nielsen, 2001), they should explain precisely what error has occurred (Nielsen, 2001), they should be placed directly to the right of where the error occurred (Seckler et al., 2012), they should include suggested alternatives (Resch & Zimmer, 2013), and they should be polite (Tzeng, 2004).

5.1.3 Some Users Translate Their Expectations from Consumer Tools to Niche Tools Like Ocean Navigator

Ocean Navigator was a completely new system to all but one participant. It thus makes sense that participants used already existing mental models to help them define their expectations and understanding of something unknown. This was likely happening much of the time, but it

was obvious when participants out loud compared Ocean Navigator to another system like Adobe, MATLAB, or Python. The main issues identified through comparison, menus, error messages, and date selection, have already been addressed previously; however, I do recommend looking at other systems to see how they solve issues Ocean Navigator is trying to solve.

5.2 LIMITATIONS

As with any research, there were instances of bias that could implicate the findings of this research. I was a Library Studies student who is particularly interested in usability and hopes to work in public libraries. In future research, it could be beneficial to involve researchers from other fields, including computer science. Careful research, planning, use of mixed methods, and consultation with thesis committee members were put into all aspects of this project to mitigate the unavoidable problem of bias. There is the potential that my inherent biases nonetheless affected the findings presented. What I found interesting, could be uninteresting to others and vice versa. I was also relatively new to qualitative research, having only been introduced to it in my graduate studies course *Research Methods* so lack of experience could still be a concern. One example of this lack of experience was that I did not think to record audio while participants were completing the SUS until multiple participants had already completed their sessions. I realized after a few sessions that participants were giving valuable feedback while completing the SUS and decided to start recording audio. However, since not all the participants were recorded while completing the SUS, valuable data was lost. I also only used the screen capture recordings for timing the tasks and clarification if it was unclear what was happening by the audio recording. The vast majority of analysis was performed on the transcriptions. There is the

potential that important data was missed that was in the screen capture recording but not clear in the audio recording. In future studies, it could be beneficial to analyze both transcriptions of audio and screen capture recordings.

While uninvolved observation of the speak aloud methodology might have been ideal, I did make suggestions at times to keep the session progressing. If a participant was completely stuck and it seemed like there would be nothing to learn from letting them continue to try, I would give just enough guidance to get the participant going again. The most common suggestion was to refresh the page. This was not something that the developers suggested but rather something I found helpful while exploring Ocean Navigator themselves. I also might not have always given the ideal suggestion. I did not have access to the recordings of the expert user until after all the participant sessions were complete and so I had not seen the ideal pathway the expert took to get the result the task description was asking for. This could have actually been beneficial in finding the kind of information we were looking for though. Since the participants were not given expert advice, they were not being taught something new about how to use the system and expert knowledge was not introduced to them throughout their session.

Unexpected events that occurred during data collection could have made a difference in what happened and thus on findings. For example, during one of the sessions, the computer that was being used ran out of battery. The pause in the session while I plugged in and restarted the computer could have impacted how the participant felt about the process and would have affected some of the timing of tasks. Other interruptions included hotel staff entering the room during a session, the computer indicating that it needed to be plugged in, the screen recording

software malfunctioning, etc. To minimize these unexpected events, a private room was booked for testing. The conference setting the research sessions took place in could also be a limitation. Attending a conference can be a tiring experience. There's so much to fit in and so little time to do it in. Many people take the opportunity to network with people in their field of study, present their research, attend sessions by other researchers they are interested in, etc. While there are benefits to participating in research, it is one more thing on top of all the things they are trying to fit in within a short period of time which can lead to fatigue that could impact the research. Future research could benefit from a more controlled, low-stress setting.

5.3 SUMMARY OF RECOMMENDATIONS

Many recommendations for immediate (or as soon as possible) modifications were presented throughout sections 5.1 and 5.2 in broad detail. Below they are summarized for convenience.

The following are general recommendations:

- Define main vocabulary and have it easily accessible through a reference tool at all times.
- Implement visual hierarchy, especially for the dataset information page.
- Use illustrations, or screen captures, with concise textual enhancers on the help page.
- Include error messages that are readable by a non-expert human, that explain precisely what error has occurred, that are placed directly to the right of where the error occurred, that include suggested alternatives, and that are polite.
- Look at other systems that are well known and highly used to see how they solve issues similar to those of Ocean Navigator.

The following are more specific recommendations:

- For date selection, use three drop down menus where a user selects the day in the first dropdown, the month in the second, and the year in the third.
- For the depth selection, use a text input field for numerical data and a radio button for the “bottom” option.
- For the draw a line function, provide feedback like status information or progress indicators.
- For entering coordinates, include a search history and state entry restrictions explicitly.
- To represent velocity, use arrows that are distinguished by colour and/or size and include explanatory notes as well as a legend.
- Do not use the same term (velocity) to represent two different concepts (arrows and contours). Change or delete one instance.
- Make the clickable area large for all clickable targets.

5.4 SUMMARY OF FUTURE RESEARCH SUGGESTIONS

Many suggestions for future research possibilities were presented throughout sections 5.1 and 5.2 in broad detail. Below they are summarized for convenience.

- Conduct interviews and observations to determine focused vocabulary that is more familiar to typical users.
- Determine if users with more experience working with this type of system gave lower scores because they had higher expectations.
- Involve researchers from other fields.

- Analyze screen capture recordings.
- Conduct sessions in a more controlled, low-stress setting.

REFERENCES

- Accot, J., & Zhai, S. (1997). Beyond Fitts' law: Models for trajectory-based HCI tasks. *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*. 295-302. <https://doi.org/10.1145/258549.258760>
- Ackerman, P., & American Psychological Association. (2011). *Cognitive fatigue: Multidisciplinary perspectives on current research and future applications* (1st ed., APA science volumes). American Psychological Association.
- Bailey, R., Barnum, C., Bosley, J., Chaparro, B., Dumas, J., Ivory, M., . . . Zimmerman, D. (2006). *Research-based web design & usability guidelines*. Dept. of Health and Human Services.
- Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An empirical evaluation of the System Usability Scale. *International Journal of Human-Computer Interaction*, 24(6), 574-594. <https://doi.org/10.1080/10447310802205776>
- Bangor, A., Kortum, P.T., and Miller, J.T. (2009). Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of Usability Studies*. 4(3) 114-123. <https://uxpajournal.org/determining-what-individual-sus-scores-mean-adding-an-adjective-rating-scale/>
- Bargas-Avila, J. A., Brenzikofer, O., Tuch, A. N., Roth, S. P., & Opwis, K. (2011). Working towards usable forms on the world wide web: optimizing date entry input fields. *Advances in Human-Computer Interaction, 2011*. <https://doi.org/10.1155/2011/202701>
- Bargas-Avila, J.A., Orsini, S., Piosczyk, H., Urwyler, D., & Opwis, K. (2011). Enhancing online forms: Use format specifications for fields with format restrictions to help respondents. *Interacting with Computers*, 23(1), 33-39. <https://doi.org/10.1016/j.intcom.2010.08.001>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101. <https://doi.org/10.1191/1478088706qp063oa>
- Brooke, J. (1996). SUS: a "quick and dirty" usability scale. In P.W.Jordan, B. Thomas, B.A. Weerdmeester, and I.L. McClelland (Eds.) *Usability Evaluation in Industry* (189-194). Taylor and Francis.
- Brooke, J. (2013). SUS: a retrospective. *Journal of usability studies*, 8(2), 29-40. <https://uxpajournal.org/sus-a-retrospective/>
- Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, 47(6), 381-391. <https://doi.org/10.1037/h0055392>

- Galitz, W. O. (2007). *The essential guide to user interface design: an introduction to GUI design principles and techniques*. John Wiley & Sons.
- Goodchild, M. (2009). NeoGeography and the nature of geographic expertise. *Journal of Location Based Services*, 3(2), 82-96. <https://doi.org/10.1080/17489720902950374>
- Hennig, S., Vogler, R., & Wasserburger, W. W. (2016). Usability and accessibility of web maps: Considering new user groups and their requirements. *International Journal of Geoinformatics*, 12(4), 17-23. <https://journals.sfu.ca/ijg/index.php/journal/article/view/987>
- Hovmöller, E. (1949). The trough-and-ridge diagram. *Tellus*. 1(2). 62-66. <https://doi.org/10.3402/tellusa.v1i2.8498>
- Johnson, J. (2014). *Designing with the mind in mind: simple guide to understanding user interface design guidelines* (2nd ed.). Elsevier.
- Kiefer, P., Giannopoulos, I., Anagnostopoulos, V. A., Schöning, J., & Raubal, M. (2017). Controllability matters: The user experience of adaptive maps. *GeoInformatica*, 21(3), 619-641. <https://doi.org/10.1007/s10707-016-0282-x>
- Kramers, E. R. (2008). Interaction with maps on the internet: A user centred design approach for the atlas of Canada. *The Cartographic Journal*, 45(2), 98-107. <https://doi.org/10.1179/174327708X305094>
- Krug, S. (2014). *Don't make me think, revisited: A common sense approach to web usability*. New Riders.
- Krug, S. (2010). *Rocket surgery made easy: The do-it-yourself guide to finding and fixing usability problems*. New Riders.
- Kubíček, P., Šašík, Č., Stachoň, Z., Štěřba, Z., Apeltauer, J., & Urbánek, T. (2017). Cartographic design and usability of visual variables for linear features. *The Cartographic Journal*, 54(1), 91-102. <https://doi.org/10.1080/00087041.2016.1168141>
- Li, A. (2017, June 11). *Dropdowns: Design guidelines*. Retrieved August 31, 2020, from <https://www.nngroup.com/articles/drop-down-menus/>
- Mahajan, R., & Shneiderman, B. (1997). Visual and textual consistency checking tools for graphical user interfaces. *IEEE Transactions on Software Engineering*, 23(11), 722-735. <https://doi.org/10.1109/32.637386>
- Merriam-Webster. (n.d.). Jargon. In *Merriam-Webster.com dictionary*. Retrieved August 26, 2020, from <https://www.merriam-webster.com/dictionary/jargon>

- Meyniel, F., Sergent, C., Rigoux, L., Daunizeau, J., & Pessiglioni, M. (2013). Neurocomputational account of how the human brain decides when to have a break. *Proceedings of the National Academy of Sciences*, 110(7), 2641-2646. <https://doi.org/10.1073/pnas.1211925110>
- Nelson, E. (2000). The impact of bivariate symbol design on task performance in a map setting. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 37(4), 61-78. <https://doi.org/10.3138/V743-K505-5510-66Q5>
- Newman, G., Zimmerman, D., Crall, A., Laituri, M., Graham, J., & Stapel, L. (2010). User-friendly web mapping: Lessons from a citizen science website. *International Journal of Geographical Information Science*, 24(12), 1851-1869. <https://doi.org/10.1080/13658816.2010.490532>
- Nielsen, J. (2001, June 23). Error Message Guidelines. Retrieved September 04, 2020, from <https://www.nngroup.com/articles/error-message-guidelines/>
- Nivala, A.-M., Sarjakoski, L., & Sarjakoski, T. (2007). Usability methods familiarity among map application developers. *International Journal of Human - Computer Studies*, 65(9), 784-795. <https://doi.org/10.1016/j.ijhcs.2007.04.002>
- Nordquist, Richard. (2020, February 11). Definition and Examples of Jargon. Retrieved from <https://www.thoughtco.com/what-is-jargon-1691202>
- Norman, D. (2013). *The design of everyday things (Revised and expanded ed.)*. Basic Books, a member of Perseus Books Group.
- Oakley, N.S., & Daudert, B. (2016). Establishing best practices to improve usefulness and usability of web interfaces providing atmospheric data. *Bulletin of the American Meteorological Society*, 97(2), 263-274. <https://doi.org/10.1175/BAMS-D-14-00121.1>
- Ocean Navigator. (n.d.). Retrieved September 30, 2020, from <https://github.com/DFO-Ocean-Navigator>
- Olson, J. M. (1981). Spectrally encoded two-variable maps. *Annals of the Association of American Geographers*, 71(2), 259-276. <https://doi.org/10.1111/j.1467-8306.1981.tb01352.x>
- Panchaud, N. H., & Hurni, L. (2018). Integrating cartographic knowledge within a geoportal: Interactions and feedback in the user Interface. *Cartographic Perspectives*, 89, 5-24. <https://doi.org/10.14714/CP89.1402>
- Play, A. o. (2019, April 12). *Maps - Navigate & Explore*. Retrieved April 2019, from <https://play.google.com/store/apps/details?id=com.google.android.apps.maps>

- Resch, B., & Zimmer, B. (2013). User experience design in professional map-based geo-portals. *User Experience Design in Professional Map-Based Geo-Portals*, 1015-1037. <https://doi.org/10.3390/ijgi2041015>
- Robertson, I., & Kortum, P. (2017). The effect of cognitive fatigue on subjective usability scores. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting 61(1)*, 1461-1465. <https://doi.org/10.1177/1541931213601850>
- Roth, R. E. (2015). interactivity and cartography: A contemporary perspective on user interface and user experience design from geospatial professionals. *Cartographica*, 94-115. <https://doi.org/10.3138/carto.50.2.2427>
- Sauro, J. (2011). *A practical guide to the System Usability Scale: Background, benchmarks & best practices*. Measuring Usability LLC.
- Seckler, M., Tuch, A., Opwis, K., & Bargas-Avila, J. (2012). User-friendly locations of error messages in web forms: Put them on the right side of the erroneous input field. *Interacting with Computers 24(3)*, 107-118. <https://doi.org/10.1016/j.intcom.2012.03.002>
- Smith, S. L., & Mosier, J. N. (1986). *Guidelines for designing user interface software*. Mitre Corporation.
- Sonstebly, A., & DeJonghe, J. (2013). Usability testing, user-centered design, and LibGuides subject guides: A case study. *Journal of Web Librarianship*, 7(1), 83-94. <https://doi.org/10.1080/19322909.2013.747366>
- Stone, D., Jarret, C., Woodroffe, M., & Minocha, S. (2005). *User interface design and evaluation*. Elsevier.
- Tzeng, J. Y. (2004). Toward a more civilized design: studying the effects of computers that apologize. *International Journal of Human-Computer Studies*, 61(3), 319-345. <https://doi.org/10/1016/j.ijhcs.2004.01.002>
- Williams, S. C., & Foster, A.K. (2011). Promise fulfilled? An EBSCO discovery service usability study. *Journal of Web Librarianship*, 5(3), 179-198. <https://doi.org/10.1080/19322909.2011.597590>
- Wikström, D. (2017). *Knowledge transfer between user interfaces: Exploring mental model influence in learning new UI - A case study at Skatteverket*. (Publication No. diva-136836) [Master's thesis, Umeå University]. DIVA Portal. <https://urn.kb.se/resolve?urn=urn%3Anbn%3Ase%3Aumu%3Adiva-136836>
- Worldview: Explore Your Dynamic Planet. (n.d.) Retrieved July 09, 2020, from <https://worldview.earthdata.nasa.gov/>

You, M., Chen, C.-w., Liu, H., & Lin, H. (2007). A usability evaluation of web map zoom and pan functions. *International Journal of Design*, 1(1), 15-25.
<http://www.ijdesign.org/index.php/IJDesign/article/view/31>

APPENDIX A Test Script

Testing the Ocean Navigator

Revised from *Rocket Surgery Made Easy* © 2010 Steve Krug

Web browser should be open to the Ocean Navigator website

Hi, my name is Ali, and I'm going to be walking you through this session today.

Before we begin, I have some information for you, and I'm going to read it to make sure that I cover everything.

You probably already have a good idea of why we asked you here, but let me go over it again briefly. We're asking people to try using the Ocean Navigator's website so we can see whether it works as intended. The session should take between 15 and 20 minutes.

The first thing I want to make clear right away is that we're testing the *site*, not you. You can't do anything wrong here. In fact, this is probably the one place today where you don't have to worry about making mistakes. Any mistakes you make provide valuable information, since they likely highlight areas within the tool that need revision.

As you use the site, I'm going to ask you to try to think out loud as much as possible: to say what you're looking at, what you're trying to do, and what you're thinking. This will be a big help to us. We're doing this to improve the site, so we need to hear your honest reactions.

If you have any questions as we go along, just ask them. I may not be able to answer them right away, since we're interested in how people do when they don't have someone sitting next to them to help. But if you still have any questions when we're done, I'll try to answer them then. If you need to take a break at any point, just let me know.

With your permission, we're going to record what happens on the screen and our conversation. The recording will only be used to help us figure out how to improve the site, and it won't be seen or heard by anyone except the people working on this project. It helps me, because I don't have to take as many notes.

If you would, I'm going to ask you to sign a consent form for us. This project has received ethics approval from the Faculty of Management, but ultimately you decide if you want to proceed. The form says that we have your permission to record you, and that the recording will only be seen and heard by the people working on the project.

Give them a consent form and a pen

Once they sign it, START the SCREEN RECORDER

OK. Before we look at the site, I'd like to ask you just a few questions.

1. Had you heard of the Ocean Navigator before being asked to participate in this study?
2. If yes, what do you know about it?

OK, great. We're done with the questions, and we can start looking at things.

Hand them the mouse

Now I'm going to ask you to try doing some specific tasks. I'm going to read each one out loud and give you a printed copy.

[tasks will be developed and finalized with Fisheries and Oceans Canada and will be basic information retrieval using the Ocean Navigator]

As much as possible, it will help us if you can try to think out loud as you go along.

- Hand the participant the first scenario and read it aloud.**
- Allow the user to proceed until you feel no value is being added to the study, or the user becomes very frustrated. Remind the user that we are testing the site, not them, as necessary.**
- Repeat for each task or until time runs out.**

Do you have any questions for me, now that we're done?

- Give them their incentive**
- Stop the screen recorder and save the file**
- Thank them and escort them out**

APPENDIX B Consent Forms

CONSENT FORM – SECOND STREAM



CONSENT FORM

Usability Testing: Ocean Navigator

You are invited to take part in a research study being conducted by me, Alison McLaughlin, a graduate student in the School of Information Management, as part of my master's degree at Dalhousie University. The purpose of this research is to evaluate the usability of a program called Ocean Navigator. We will be asking you to complete tasks on the Ocean Navigator tool and to complete a survey on how usable they find the program. The study is funded by Marine Environmental Observation Prediction and Response Network. I will write up the results of this research in my Masters dissertation and academic papers, and may share my results more widely in public presentations.

As a participant in the research you will be asked to complete tasks on a web application called Ocean Navigator and asked to answer 12 questions in a survey conducted over the internet using Opinio software. All responses will be saved on a secure Dalhousie server and processed using R Statistical software. The survey does not ask for your name.

Your participation in this research is entirely your choice. You do not have to answer questions that you do not want to answer, and you are welcome to stop the survey at any time if you no longer want to participate. All you need to do is close your browser. I will not include any incomplete surveys in my analyses. However, if you do complete your survey, if you change your mind later, I will not be able to remove the information you provided because the surveys are completed anonymously, so I would not know which one is yours.

Information that you provide to me will be collected anonymously, which means that there will be no questions asked in the survey that asks for identifying details such as your name or email address. Only my supervisor and I will have access to the survey results. I will describe and share general findings in several ways including presentations, my thesis, journals, etc. I will destroy all information 12 months after reporting the results.

The risks associated with this study are no greater than those you encounter in your everyday life.

There will be no direct benefit to you in participating in this research and you will not receive compensation. The research, however, might contribute to new knowledge on the usability of Ocean Navigator. If you would like to see how your information is used, please feel free to contact us for a copy of the final report. The final report will also be made publicly available. You should discuss any questions you have about this study with me (Alison McLaughlin). Please ask as many questions as you like. If you have questions later, please feel

free to contact me or my supervisor. My email is afmclaughlin@dal.ca and mike.smit@dal.ca is my supervisor's.

If you have any ethical concerns about your participation in this research, you may contact Research Ethics, Dalhousie University at (902) 494-1462, or email ethics@dal.ca (and reference REB file # 20XX-XXXX).”

If you agree to complete the survey, please follow the link here.

CONSENT FORM

Project title: Usability Testing: Ocean Navigator

Lead researcher: Alison McLaughlin, School of Information Management afmclaughlin@dal.ca

Other researchers

Dr. Mike Smit, School of Information Management mike.smit@dal.ca

Funding provided by: Marine Environmental Observation Prediction and Response Network

Introduction

You are invited to take part in a research study being conducted by me, Alison McLaughlin, a graduate student, at Dalhousie University as part of my Master of Information degree program. Choosing whether or not to take part in this research is entirely your choice. The information below tells you about what is involved in the research, what you will be asked to do and about any benefit, risk, inconvenience or discomfort that you might experience.

You should discuss any questions you have about this study with me (Alison McLaughlin). Please ask as many questions as you like. If you have questions later, please do not hesitate to contact me then.

Purpose and Outline of the Research Study

The purpose of this research is to evaluate the usability of a program called Ocean Navigator. Usability testing is intended to identify areas for improvement by testing how satisfied the user is with navigating a digital environment. Approximately 10 participants will complete tasks while thinking out loud

Who Can Take Part in the Research Study

You may participate in this study if you consider yourself to have at least some experience working with ocean data. You may not participate in this study if you have previous experience using Ocean Navigator.

[Field]What You Will Be Asked to Do

You will be asked to complete a variety of information-seeking tasks on the Ocean Navigator, and to speak aloud while doing so. Following the tasks, you will be asked to complete a short survey. Your interactions on the screen will be recorded (but no video of you will be captured), as well as audio recording of your speaking aloud and question responses. We expect the study to take no longer than one hour for you to complete.

Possible Benefits, Risks and Discomforts

Participating in the study might not benefit you directly, but we might learn things that will benefit others. Ocean Navigator is not a tool that is widely known, so participation in this study will expose oceanographers to a potentially useful tool for their research. Indirectly, potential future users of the Ocean Navigator will benefit from a tool that better meets their needs. The risks associated with this study are minimal, and there are no known risks for participating in this research beyond being bored or fatigued. However, you will be offered breaks between activities to reduce these risks.

Compensation / Reimbursement

As a small token of thanks for your time, we will give you \$10 cash following the session.

How your information will be protected:

Participants will not be anonymous to the researchers, but their data will be anonymized in all reports/presentations etc. Therefore, it will not be possible for members of the public to trace information/results to any one participant. The researcher will store your data without personal identifiers, though because of a) the small sample size and b) the audio recording for in-person participants, even demographic data may be sufficient to re-identify participants. The data will therefore be stored confidentially in approved Dalhousie file-sharing systems, on password-protected computers, or on portable storage that is locked in a cabinet when not in use.

Information that you provide to us will be kept private. Only the research team at Dalhousie University will have access to this information. We will describe and share our findings in my thesis. ***You will not be identified in any way in our reports.*** The people who work with us have an obligation to keep all research information private. Also, we will use a participant number (not your name) in our written and computer records so that the information we have about you contains no names. All your identifying information will be securely stored.

If You Decide to Stop Participating

You are free to leave the study at any time. If you decide to stop participating at any point in the study, you can also decide whether you want any of the information that you have contributed up to that point to be removed or if you will allow us to use that information. After that time, it will become impossible for us to remove it because it will already be anonymized.

How to Obtain Results

You will be invited to contact us for a copy of the final report. The final report will also be made publicly available.

Questions

We are happy to talk with you about any questions or concerns you may have about your participation in this research study. Please contact Alison McLaughlin (at afmclaughlin@dal.ca) [or Dr. Mike Smit (at mike.smit@dal.ca)] at any time with questions, comments, or concerns about the research study. We will also tell you if any new information comes up that could affect your decision to participate.

If you have any ethical concerns about your participation in this research, you may also contact Research Ethics, Dalhousie University at (902) 494-1462, or email: ethics@dal.ca (and reference REB file # 20XX-XXXX).

Signature Page

Project Title: Usability Testing: Ocean Navigator

Lead Researcher: Alison McLaughlin, School of Information Management afmclaughlin@dal.ca

I have read the explanation about this study. I have been given the opportunity to discuss it and my questions have been answered to my satisfaction. I understand that I have been asked to take part in a usability study that will occur at a location acceptable to me, and that the session will be recorded. I understand direct quotes of things I say may be used without identifying me. I agree to take part in this study. My participation is voluntary and I understand that I am free to withdraw from the study at any time, until I have left my session.

NameSignatureDate

APPENDIX C Recruitment

INDIRECT RECRUITMENT – FIRST STREAM RECRUITMENT EMAIL

Subject: Invitation: Ocean Navigator User Experience Study

Dear <recipient>,

<Where applicable, a personal greeting>

We are working on a project with Fisheries and Oceans Canada to better understand the experience of new users of the Ocean Navigator, a tool for exploring ocean data.

If you would be willing to distribute the email below or put us in contact with the right people to distribute it, we would greatly appreciate it!

Thanks,

<sender: Alison or Mike>

Hello,

We're trying to make ocean model data easier to use, and as a token of appreciation for your help we're offering a < [\$10 gift card to Sobeys] [a 10\$ donation to The Ocean Conservancy]> for participating in our study.

My name is Alison McLaughlin. I am a graduate student researcher in the School of Information Management at Dalhousie University, and this is part of my graduate course work supervised by Dr. Mike Smit, a professor in my School. We are working with Fisheries and Oceans Canada on a tool called Ocean Navigator, and we need your help to test it.

We are looking for people who have no experience using the Ocean Navigator but do have experience with oceanographic data to get an idea of how user friendly the tool is to new users.

This is an opportunity to share your perspectives on the Ocean Navigator tool and its usefulness to you. You will remain anonymous in any reports we produce, though the reports and the results will be made public. Participation is entirely voluntary, and you can withdraw at any point.

If you are interested, please reply to this email indicating such. If you think somebody you know would also be a good candidate for this study, please feel free to forward them this email.

If you have any questions or concerns, please contact Alison McLaughlin or Dr. Mike Smit.

Thank you,

Alison McLaughlin
Dalhousie University
Email: afmclaughlin@dal.ca

Dr. Mike Smit
Dalhousie University
Email: mike.smit@dal.ca
Phone: (902) 494-1901
Fax: (902) 494-2451

DIRECT RECRUITMENT – FIRST STREAM
RECRUITMENT EMAIL

Subject: Invitation: Ocean Navigator Use Experience Study

Hello,

We're trying to make ocean model data easier to use, and as a token of appreciation for your help we're offering a < [\$10 gift card to Sobeys] [a 10\$ donation to The Ocean Conservancy]> for participating in our study.

My name is Alison McLaughlin. I am a graduate student researcher in the School of Information Management at Dalhousie University, and this is part of my graduate course work supervised by Dr. Mike Smit, a professor in my School. We are working with Fisheries and Oceans Canada on a tool called Ocean Navigator, and we need your help to test it.

We are looking for people who have no experience using the Ocean Navigator but do have experience with oceanographic data to get an idea of how user friendly the tool is to new users.

This is an opportunity to share your perspectives on the Ocean Navigator tool and its usefulness to you. You will remain anonymous in any reports we produce, though the reports and the results will be made public. Participation is entirely voluntary, and you can withdraw at any point.

If you are interested, please reply to this email indicating such. If you think somebody you know would also be a good candidate for this study, please feel free to forward them this email.

If you have any questions or concerns, please contact Alison McLaughlin or Dr. Mike Smit.

Thank you,

Alison McLaughlin
Dalhousie University
Email: afmclaughlin@dal.ca

Dr. Mike Smit

Dalhousie University
Email: mike.smit@dal.ca
Phone: (902) 494-1901
Fax: (902) 494-2451

MEOPEER RECRUITMENT EMAIL – FIRST STREAM

Subject: Invitation: Ocean Navigator Use Experience Study

Hey MEOPeers,

I'm Ali, and I'm a graduate student researcher in the School of Information Management at Dalhousie University supervised by Dr. Mike Smit. I'm currently working on a project where we're trying to make ocean model data easier to use. Fisheries and Oceans Canada created a tool called Ocean Navigator, and we need your help to test it. If you're willing to participate, as a token of appreciation for your help, we're offering a < [\$10 gift card to Sobeys] [a 10\$ donation to The Ocean Conservancy]> for participating in our study!

We are looking for people who have no experience using the Ocean Navigator but do have experience with oceanographic data to get an idea of how user friendly the tool is to new users.

This is an opportunity to share your perspectives on the Ocean Navigator tool and its usefulness to you. You will remain anonymous in any reports we produce, though the reports and the results will be made public. Participation is entirely voluntary, and you can withdraw at any point.

If you are interested, please reply to this email indicating such. If you think somebody you know would also be a good candidate for this study, please feel free to forward them this email.

If you have any questions or concerns, please contact me or Dr. Mike Smit.

Thank you,

Alison McLaughlin
Dalhousie University
Email: afmclaughlin@dal.ca

Dr. Mike Smit
Dalhousie University
Email: mike.smit@dal.ca
Phone: (902) 494-1901
Fax: (902) 494-2451

MEOPEER RECRUITMENT FACEBOOK POST – FIRST STREAM

Subject: Invitation: Ocean Navigator User Experience Study

Hey MEOPeers,

I'm Ali, and I'm a graduate student researcher in the School of Information Management at Dalhousie University supervised by Dr. Mike Smit. I'm currently working on a project where we're trying to make ocean model data easier to use. Fisheries and Oceans Canada created a tool called Ocean Navigator, and we need your help to test it. If you're willing to participate, as a token of appreciation for your help, we're offering a < [\$10 gift card to Sobeys] [a 10\$ donation to The Ocean Conservancy]> for participating in our study!

We are looking for people who have no experience using the Ocean Navigator but do have experience with oceanographic data to get an idea of how user friendly the tool is to new users.

This is an opportunity to share your perspectives on the Ocean Navigator tool and its usefulness to you. You will remain anonymous in any reports we produce, though the reports and the results will be made public. Participation is entirely voluntary, and you can withdraw at any point.

If you are interested, email me indicating such. If you think somebody you know would also be a good candidate for this study, please feel free to forward them this information.

If you have any questions or concerns, please contact me or Dr. Mike Smit.

Thank you,

Alison McLaughlin
Dalhousie University
Email: afmclaughlin@dal.ca

Dr. Mike Smit
Dalhousie University
Email: mike.smit@dal.ca
Phone: (902) 494-1901
Fax: (902) 494-2451

APPENDIX D Task Descriptions

1. Please imagine that your project requires that you know something about the water conditions in various areas of the Atlantic Ocean at various times, and that you know the GIOPS model might have this data, accessible through the Ocean Navigator. You know that there are Ocean Navigator features built around working with a single point, a line, or an area on the map. You also know that the GIOPS model has projected and historic water temperature
 - a. Access the daily average water temperature at various depths (a water temperature profile) at 44.2135, -63.1242 on the Scotian Shelf for April 10, 2016. Once you have created the profile, save it as a png file.
 - b. You're interested in the water conditions likely to be experienced at a particular site from April 2016 to January 2018. Ideally you would have had a buoy (mooring) physically installed, but it's too late for that now! Use this tool to access the monthly average water temperature at the bottom of the Atlantic Ocean at 42.7088, -60.9059.
 - c. Create a map of a specific area (use the latitude and longitude points below) that shows both the average surface temperature and the average ocean current (i.e. water velocity) for the month of September 2018.

Latitude	Longitude
36.52	-42.77
36.52	-39.99
35.05	-39.99
35.05	-42.77

- d. Zoom in to the north-east coast of Labrador. Draw a line from the coast of Labrador to past the edge of the Labrador Shelf. Use a Hovmöller Diagram to show the average daily sea ice extent for the last week of January 2019 (i.e., a diagram to show how far offshore the ice concentration is more than 15%?).
2. Ocean models can be assessed by comparing measurements actually collected on site (observed in situ) to the values predicted by the model. There is a global standard for doing this type of comparison, commonly referred to as calculating “class 4 metrics”. Imagine that you are interested in assessing the accuracy of the GIOPS Daily model output for May 1, 2019, with a particular interest for ocean conditions near Labrador.
 - a. Visually compare temperature profiles as observed to those modeled.
 - b. The same comparison but include Climatology data in the comparison.
 - c. Now compare the observed data to the FOAM model
3. You have historic biomass abundances from April 2004 and October 2004 survey data. The study area is defined by the CSV labelled “Total Area.csv”. You are interested in understanding the ocean conditions during the survey period, and so are looking for climatology data for the area and specifically the GLORYS v4 Climatology dataset.
 - a. Create and download a png of the bottom temperature climatology. Include water velocity arrows on the plot. Remove bathymetry contours.
 - b. Download an R script to generate this same plot in R.

APPENDIX E System Usability Scale

In the following scales, please circle the number that best represents your feelings toward the Ocean Navigator program.

How unlikely (1) or likely (5) would you be to recommend this program to a fellow researcher in your field?

1	2	3	4	5
Not at all likely		Neutral		Extremely likely

Please rate your agreement with the following statements:

“I understand what I can do here and who this site is for.”

1	2	3	4	5
Strongly disagree		Neutral		Strongly agree

“I think that I would like to use this system frequently

1	2	3	4	5
Strongly disagree		Neutral		Strongly agree

“I found the system unnecessarily complex”

1	2	3	4	5
Strongly disagree		Neutral		Strongly agree

“I thought the system was easy to use”

1	2	3	4	5
Strongly disagree		Neutral		Strongly agree

“I think that I would need the support of a technical person to be able to use this system”

1	2	3	4	5
Strongly disagree		Neutral		Strongly agree

“I found the various functions in this system were well integrated”

1	2	3	4	5
Strongly disagree		Neutral		Strongly agree

“I thought there was too much inconsistency in this system”

1	2	3	4	5
Strongly disagree		Neutral		Strongly agree

“I would imagine that most people would learn to use this system very quickly”

1	2	3	4	5
Strongly disagree		Neutral		Strongly agree

“I found the system very cumbersome to use”

1	2	3	4	5
Strongly disagree		Neutral		Strongly agree

“I felt very confident using the system”

1	2	3	4	5
Strongly disagree		Neutral		Strongly agree

“I needed to learn a lot of things before I could get going with this system”

1	2	3	4	5
Strongly disagree		Neutral		Strongly agree

APPENDIX F Initial Codes

- COMPARISON
 - Adobe
 - MATLAB
 - other
 - Python
- EMOTIONS
 - Conflicted
 - Confusion
 - Constrained
 - Frustration
 - satisfaction
- SUGGESTION
 - Area clickable
 - Bottom
 - Date (select)
 - Error message
- TASK STATUS
 - Backtracking
 - Complete
 - Incomplete

APPENDIX G Codes Discovered During Phase One of Coding

- COMPARISON
 - worldview.earthdata.nasa.gov/
- EMOTIONS
 - Annoyed
 - Curiosity
 - Defeat
 - Disbelief
 - Nervous
 - Uncertainty
 - unsure
- SUGGESTIONS
 - Area more shapes
 - Arrow spacing
 - Date (start end)
 - Explore
 - Export products
 - Feedback
 - Help
 - Memory
 - Menu
 - Plot title
 - Simple language
 - Surface
 - tutorial
- TO BE CLASSIFIED
 - Asked for help
 - Blaming server
 - Blaming something other than the system
 - Blaming themselves instead of the system
 - Blaming wifi
 - Compare datasets
 - Dataset information
 - Failed attempt?
 - Inconsistency
 - Inefficient
 - Justifying the system not working
 - Misunderstood line
 - Needed help
 - Not sure if complete
 - Participant has seen ocean navigator before
 - Refresh
 - Something took longer than expected
 - Thought single click would end line

APPENDIX H Renaming of Codes for Phase Two

- **COMPARISON**
 - Adobe
 - MATLAB
 - other
 - Python
 - worldview.earthdata.nasa.gov/
- **EMOTIONS**
 - Conflicted
 - Confusion
 - Constrained
 - Frustration
 - Satisfaction
 - Annoyed --> merged with frustration bc similar definitions
 - Curiosity
 - Defeat
 - (deleted) Disbelief
 - Nervous
 - Uncertainty
 - Unsure
- **SUGGESTION**
 - Area clickable --> renamed PROBLEM: clicking area is difficult
 - Bottom --> separated in SUGGESTION: bottom and PROBLEM: bottom not obvious
 - Date (select) --> separated in SUGGESTION: date (select) and PROBLEM: date hard to select
 - Error message --> separated in SUGGESTION: error message and PROBLEM: lack of error message
 - Area more shapes --> renamed PROBLEM: area selection limiting (only lines)
 - Arrow spacing --> PROBLEM: arrow size/spacing and SUGGESTION: arrow size/spacing
 - Date (start end) --> PROBLEM: date (start end)
 - Explore --> PROBLEM: explore
 - Export products
 - Feedback
 - Help --> PROBLEM: no help page and dataset information
 - Memory --> PROBLEM: lack of memory and SUGGESTION: memory
 - Menu
 - Plot title
 - Simple language
 - Surface --> renamed PROBLEM: surface not obvious
 - Tutorial
- **TO BE CLASSIFIED**
 - Asked for help --> changed to HELP: asked for
 - Blaming server --> renamed BLAME: server
 - Blaming something other than the system --> renamed BLAME: other

- Blaming themselves instead of the system--> renamed BLAME: self
- Blaming wifi --> renamed BLAME: wifi
- Compare datasets --> merged with PROBLEM: no help page
- Dataset information--> renamed and split into PROBLEM: no dataset information and SUGGESTION: dataset information
- Failed attempt? --> deleted
- Inconsistency--> changed to SUGGESTION: map pop up change
- Inefficient--> kept
- Justifying the system not working--> changed to BLAME: other
- Misunderstood line--> changed to PROBLEM: line not friendly
- Needed help--> HELP: researcher provided
- Not sure if complete--> deleted
- Participant has seen ocean navigator before--> kept
- Refresh--> changed to participant had to refresh
- Something took longer than expected--> PROBLEM: length of task
- Thought single click would end line--> PROBLEM: line function unfriendly

TASK STATUS --> removed

APPENDIX I Codes Used for Phase Two

- BLAME
 - Other
 - Self
 - Server
 - wifi
- COMPARISON
 - Adobe
 - MATLAB
 - Other
 - Python
 - worldview.earthdata.nasa.gov/
- EMOTIONS
 - Conflicted
 - Confusion
 - Constrained
 - Curiosity
 - Defeat
 - Frustration
 - Nervous
 - Satisfaction
 - Uncertainty
 - unsure
- HELP
 - Asked for
 - Researcher provided
- PROBLEM
 - Area selection limiting (only lines)
 - Arrow size/spacing
 - Bottom not obvious
 - Clicking area is difficult
 - Dataset information unfriendly
 - Date (start end)
 - Date hard to select
 - Explore
 - Jargon
 - Lack of error message
 - Lack of memory
 - Length of task
 - Line function unfriendly
 - No help page
 - Selecting area difficult
 - surface
- SUGGESTION
 - Arrow size/spacing
 - Bottom

- Dataset information
- Date (select)
- Error message
- Export products
- Feedback
- Map pop up change
- Memory
- Menu
- Plot title
- Simple language
- tutorial
- MISC
 - Inefficient
 - Participant had to refresh
 - Participant has seen ocean navigator before

APPENDIX J Final List of Codes with Definitions

“BLAME” was used when the participant blamed an issue on anything other than the system (Ocean Navigator) itself or justified the system not working because of something other than the system

“BLAME: other” was used when the participant placed blame on something other than the system but on something that was not quite tangible. Participants justified issues by saying “but that’s not unique to this program”, “I understand it’s a lot of data”, and “but that would take more coding”.

“BLAME: self” was used when the participant placed blame on themselves rather than the system. Participants justified issues by saying things like “maybe there’s an easy way and I just don’t see it”, “I’m just apparently not very observant”, “maybe I’m just getting tired”, or “I wouldn’t say I felt confident using the system but that’s more an artifact of my background than the system”.

“BLAME: server” was used when the participant placed blame on the server rather than the system. Participants justified issues by saying “there must be a lot of people using Ocean Navigator right now” and “because it seems like it’s almost like a server issue”.

“BLAME: wifi” was used when the participant placed blame on the wifi rather than the system. Participants justified issues by saying “let’s blame the wifi of the hotel” and “I continue to be confused about whether is this data available or is my internet slow”.

“COMPARISON” was used when the participant compared Ocean Navigator to another system.

“COMPARISON: Adobe” was used when the participant compared Ocean Navigator to Adobe. The participant who used this comparison said “ but maybe it’s my Adobe trained brain”

“COMPARISON: MATLAB” was used when the participant compared Ocean Navigator to MATLAB. Participants comparing to MATLAB said “least favorite features about old MATLAB inaccurate error messages”, “if I was making this in MATLAB myself I would then draw a contour where the ice concentration is or the ice concentration fraction is 0.15”, and “I feel like I don’t know I guess because I’m a quantitative person and I work with things like MATLAB Python or whatever I want to be able to export that data as like numbers”

“COMPARISON: other” was used when the participant compared Ocean Navigator to something intangible. The participant who used this comparison said “yeah I mean that’s you can do it pretty easily it’s not like it’s that hard but I feel like these days we’re so used to just like well of course it knows what I want to do I want the next thing right there I don’t want to have to select another tool or something like that”

“COMPARISON: Python” was used when the participant compared Ocean Navigator to Python. The participant who used this comparison said “I feel like I don’t know I guess because I’m a quantitative person and I work with things like MATLAB Python or whatever I want to be able to export that data as like numbers”

“COMPARISON: worldview.earthdata.nasa.gov/” was used when the participant compared Ocean Navigator to worldview.earthdata.nasa.gov/. The participant who used this comparison demonstrated features they liked on worldview.earthdata.nasa.gov/ and expressed that they would be helpful features to incorporate into Ocean Navigator

“EMOTIONS” was used when a participant expressed an emotion either through directly saying they were feeling it or implying that they were feeling it

“EMOTIONS: conflicted” was used when a participant said “I feel conflicted”

“EMOTIONS: confusion” was used when a participant either said “I’m confused” or expressed confusion through saying something like “why is it undefined”, “I don’t know why...”, “I wonder if...”, “so maybe there’s something wrong again”, “I don’t know how...”, “but which one”, “I don’t know what...”, “what’s the difference between...”, “how do I...”, “I don’t know if...”, “what I don’t know is...”, or “I don’t understand”.

“EMOTIONS: constrained” was used when a participant either said “I feel constrained” or expressed the feeling by saying “now I can’t do anything”

“EMOTIONS: curiosity” was used when a participant either said “I’m curious” or expressed curiosity through saying “what if I...”.

“EMOTIONS: defeat” was used when a participant expressed a feeling of defeat by saying something like “I think I’m just gonna call it a day”, “I’m gonna go on to the next one” (without completing the task), “I failed totally”, “there’s a way to do it, but I have no idea”, or “I don’t know what to do”

“EMOTIONS: frustration” was used when a participant either said something like “I feel frustrated” or expressed frustration through saying something like “I broke this thing”, “every time that I have to include something I want to die”, “it won’t stop yelling at me”, or “oh now it’s upset at me”.

“EMOTIONS: nervous” was used when a participant either said something like “I’m nervous/worried/scared”.

“EMOTIONS: satisfaction” was used when a participant expressed a feeling of satisfaction by saying something like “perfect”, “that’s nice”, “okay great, so that’s done”, “great”, “okay” (after completing something successfully), “I’m happy with my results”, “I think I’m done with...”, “there we go”, “okay, success”, “that’s good”, “that is what I’m looking for”, or “sweet, I win”.

“EMOTIONS: uncertainty” was used when a participant was unsure if they completed the task or not and said something like “okay, am I done?” or “I think I did that right”

“EMOTIONS: unsure” was used when a participant said “I’m not sure” or expressing that they were unsure by saying “I think that means...”.

“HELP” was used when the participant asked for help and/or the researcher provided help

“HELP: asked for” was used when the participant asked for help. Participants who this code was applied to said things like “hmm any suggestions how I should do” and “is there a better way to do it than this?”.

“HELP: researcher provided” was used when the researcher provided help. This code was applied when the researcher said things like “you’re missing the date”, “so just a hint you’re trying to enter a coordinate so maybe try to find somewhere where you can enter a coordinate”, and “if you get completely stuck the easiest thing to do is just refresh the entire page”. Help involved:

- recommending a participant refresh the screen
- reminding a participant to enter a date they had not entered or one they had entered incorrectly
- suggesting a participant find somewhere to enter a coordinate
- redirecting a participant from the maximum and minimum fields to the field where they could enter a coordinate
- recommending a participant change the depth they selected
- recommending a participant change the dataset they were using

- recommending a participant reread the task description and pointing out an important piece of information as they did so
- recommending a participant focus on a more important part of a task description when they were focused on something unimportant
- showing a participant where the “virtual mooring” option was
- suggesting a participant select the surface option instead of the bottom option
- suggesting a participant change a Hovmöller diagram to show 15% sea ice extent as opposed to 14%
- giving some kind of hint related to finding the Class 4 option
- selecting an area for a participant that they had uploaded from a CSV
- showing a participant the bottom option.

“MISC” was used when something was deemed noteworthy but did not fit into a higher-level code.

“MISC: inefficient” was used when a participant described the system as inefficient. The participant who this code was applied to said the phrases “first of all, A. not efficient, B. this is also going to be extremely inefficient probably very shortly” and “ugh this should be more efficient”.

“MISC: participant had to refresh” was used when the participant refreshed the page due to it freezing.

“MISC: participant has seen ocean navigator before” was used for the one participant who recognized Ocean Navigator as something they had seen before.

“PROBLEM” was used when the participant wanted to do something and could not

“PROBLEM: area selection limiting (only lines)” was used when a participant said “is there a way to do a circle or a non-square polygon”

“PROBLEM: arrow size/spacing” was used when a participant said “I am looking for something to add just the number of arrows or the space between each arrow”

“PROBLEM: bottom not obvious” was used when a participant wanted to select the depth “bottom” but could not find the option. Participants said things like “we don't know what the bottom is”, “is there a way to know exactly what the bottom is”, and “on the list of depth options I just picked the last one I'm gonna see if that works”.

“PROBLEM: clicking area is difficult” was used when participants expressed dislike of the fact that you have to click the outline of an area instead of in the middle of an area. Participants said things like “that's also somewhat confusing sometimes because it doesn't always work if you click like that and don't you have to be like right on the line or if you're zoomed out” and “okay! I just needed to click on it there we go”

“PROBLEM: dataset information unfriendly” was used when participants wanted information on the dataset and either didn't notice that it was available or found the information difficult to comprehend. Participants said things like “maybe because some of these datasets I have heard of but some of them I have not maybe it is common knowledge to most oceanographers but because maybe I'm a bit more new to this maybe if there was a page where you could have a little bit of a description of the datasets”, “I guess I could read this thing but it's long and I don't want to read it all”, “I know nothing about those observations like where did they come from what is their source so I would be curious about that”

“PROBLEM: date (start end)” was used when the participant had difficulty selecting a range of dates, usually because it will not let you choose a start date that does not work with the end date so you have to know which one to change first.

“PROBLEM: date hard to select” was used when a participant had trouble selecting the date and had to click back through a lot of dates before getting to the one they were looking for.

“PROBLEM: jargon” was used when a participant expressed that they were unfamiliar with a term. Participants said things like “I don’t know what a Hövmoller diagram is”, “I just think it’s jargon”, “there’s an assumption that you know what class 4 means”, and “I don’t know that term”.

“PROBLEM: lack of error message” was used when the system froze instead of producing an error message

“PROBLEM: lack of memory” was used when the participant expressed unhappiness with the system not remembering a point they had just put in when they reopened the window to enter a point. Participants said “I’m scared to hit this because I’m worried I’ll lose the place I am”, “I want to enter the coordinates I got to do this again I feel frustrated... three five maybe it could remember where I’ve searched before”, and “yes, yes things that remember make my life easier”.

“PROBLEM: length of task” was used when a participant expressed that they thought the task was taking long.

“PROBLEM: line function unfriendly” was used when a participant didn’t realize that they had to double click to end a line.

“PROBLEM: no help page” was used when a participant selected a help option and there was no help.

“PROBLEM: selecting area difficult” was used when a participant tried to click and drag to create an area instead of clicking multiple points.

“PROBLEM: surface” was used when a participant thought “0m” meant zero meters from the bottom (it means surface)

“SUGGESTION” was used when a participant suggested something they would like changed about Ocean Navigator. Participants said things like “it would be nice if you could...”, “that could be something to add”, and “it would be helpful if you could...”

“SUGGESTION: arrow size/spacing” was used when the participant suggested being able to change the size or spacing of velocity arrows.

“SUGGESTION: bottom” was used when the participant suggested having the “bottom” depth option at the bottom of the list.

“SUGGESTION: dataset information” was used when the suggested for there to be information available about the datasets.

“SUGGESTION: date (select)” was used when a participant suggested that there should be an easier way to select a date. Some recommending adding an ability to type in a date rather than only clicking.

“SUGGESTION: error message” was used when a participant suggested that there should be an error message.

“SUGGESTION: export data” was used when a participant suggested that there should be an option to export the raw data.

“SUGGESTION: feedback” was used when a participant suggested that there should be “feedback” or “guidance” from the system.

“SUGGESTION: map pop up change” was used when a participant suggested that prefer the graph should pop up next to the map instead of pulling up a new map with the graph

“SUGGESTION: memory” was used when a participant suggested that the system should remember the points they had already typed in for reuse in the future.

“SUGGESTION: menu” was used when a participant suggested changes to how the menus are set up.

“SUGGESTION: plot title” was used when the participant expressed a desire to change aspects about a plot, specifically the title.

“SUGGESTION: tutorial” was used when a participant expressed a desire for a tutorial to be provided.