Website Usability Differences between Males and Females: An Eye-Tracking Evaluation of a Climate Decision Support System

LINDSAY C. MAUDLIN[a]

Department of Marine, Earth, and Atmospheric Sciences, North Carolina State University, Raleigh, North Carolina

KAREN S. MCNEAL

Department of Geosciences, Auburn University, Auburn, Alabama

HEATHER DINON-ALDRIDGE AND COREY DAVIS

State Climate Office of North Carolina, North Carolina State University, Raleigh, North Carolina

RYAN BOYLES

Southeast Climate Adaptation Science Center, U.S. Geological Survey, Raleigh, North Carolina

RACHEL M. ATKINS

Department of Marine, Earth, and Atmospheric Sciences, North Carolina State University, Raleigh, North Carolina

ABSTRACT

Decision support systems—collections of related information located in a central place to be used for decision-making—can be used as platforms from which climate information can be shared with decision-makers. Unfortunately, these tools are not often evaluated, meaning developers do not know how useful or usable their products are. In this study, a web-based climate decision support system (DSS) for foresters in the southeastern United States was evaluated by using eye-tracking technology. The initial study design was exploratory and focused on assessing usability concerns within the website. Results showed differences between male and female forestry experts in their eye-tracking behavior and in their success with completing tasks and answering questions related to the climate information presented in the DSS. A follow-up study, using undergraduate students from a large university in the southeastern United States, aimed to determine whether similar gender differences existed and could be detected and, if so, whether the cause(s) could be determined. The second evaluation, similar to the first, showed that males and females focused their attention on different aspects of the website; males focused more on the maps depicting climate information while females focused more on other aspects of the website (e.g., text, search bars, and color bars). DSS developers should consider the possibility of gender differences when designing a web-based DSS and include website features that draw user attention to important DSS elements to effectively support various populations of users.

1. Introduction

In the face of global climate change, the need to disseminate accurate climate information to decision-makers for adaptation purposes has grown (Salazar 2010); however, this is no easy feat. Consideration must be given to the climate information provided to decision-makers such that it is understandable and used appropriately (Snover et al. 2013). Further consideration should be given to the presentation of the climate information from the visualization and design of the content to the tailoring of information for the intended audience (Daron et al. 2015; Harold et al. 2016).

One way to disseminate tailored climate information to decision-makers is through a decision support system

[a] Current affiliation: Department of Geosciences, Auburn University, Auburn, Alabama.

Corresponding author: Lindsay C. Maudlin, lcm0046@auburn.edu

(DSS), a collection of information located in a central place that decision-makers can access and consult before making a decision. DSSs and similar products are widely used in many fields, including water resource management (Andreu et al. 1996), stock forecasting (Feuerriegel and Gordon 2018), and humanitarian relief management (Sahebjamnia et al. 2017). As DSSs become increasingly common, they are rarely rigorously evaluated for usability and user understanding and therefore do not always meet the needs of the decision-makers using them (Pyke et al. 2007). To ensure that this strategy is effective and that decision-makers understand the information they will ultimately use to make potentially high-impact decisions, the efficacy of these DSSs needs to be investigated through evaluations (Perry et al. 2016). An important component of a DSS evaluation is exploring how end-user characteristics, such as gender, influence the overall usability. When differences are found across groups of users, such as males versus females, they highlight areas where the DSS can be updated to be more equitable for all users.

a. PINEMAP DSS

One such DSS has been developed and evaluated. As part of the Pine Integrated Network: Education, Mitigation, and Adaptation Project (PINEMAP; http://climate.ncsu.edu/pinemap/), a USDA National Institute of Food and Agriculture project that focused primarily on loblolly pine forests in the southeastern United States, the PINEMAP DSS was developed. The web-based PINEMAP DSS provides historical and future climate and regional productivity model data specifically for decision-making foresters in this region through a suite of tools that explore parameters such as temperature and precipitation and how those parameters impact long-term seedling health. After initial PINEMAP DSS development, a beta-testing period was implemented before this evaluation.

b. Gender differences in website usability

Of special interest in this evaluation are the differences in overall usability of the PINEMAP DSS between males and females. The existing literature on gender differences within website design and usability studies is limited. Moreover, existing studies are heavily focused on e-commerce and online shopping applications and the aesthetics of these types of websites; however, these differences can shed some light on design preferences between the genders that can still be relevant to the PINEMAP DSS evaluation in this study. Cyr and Bonanni (2005) show that women are consistently less pleased with a website than men using a Likert scale rating system. In addition, this study indicates

that more men than women approve of the website's organization and layout and are able to easily navigate the website. Other studies have reported gender differences in design preferences including shapes and colors and in opinions of what is eye-catching and attractive (Moss et al. 2006; Djamasbi et al. 2007). One possible explanation for these gender differences is the manner in which males and females process information. Males tend to be selective with the information that they attempt to comprehend while females tend to attempt to comprehend all of the available information, so females are more likely to be less satisfied with a website if it does not include adequate information or quickly draw their attention to it (Simon 2001).

If users are expected to find information and be able to recall and apply that information, then a second factor, search behavior, must be considered as this can impact learning gains (Roy et al. 2003; Roy and Chi 2003). For example, important design criteria for males include minimal action, meaning less clicking or typing to navigate the website, and flexibility, meaning tasks can be completed in different ways. In contrast, females prefer learnability, which entails having a similar layout and navigation process throughout the website, and user guidance, which entails feedback and hints (Lin and Hsieh 2016).

The male-dominated IT industry could potentially be a stumbling block for women using websites mainly designed by males (Moss et al. 2006), in part because previous research has shown that gender preferences of websites and their elements tend to align best when the same gender designed the site (Moss and Gunn 2007). Boys have been shown to have higher aesthetic preferences for existing children's websites than girls, in part because of visual complexity since boys preferred higher design complexity than girls (Wang 2014). Less studied is the affective aspect of website design and how it might influence website usability. A user's affective response to a new website has been shown to influence his or her perceived usefulness and ease of use (Zhang and Li 2005). Statistically significant gender differences justify customizing a website for its audience (Simon 2001), and if a website is not customized for female viewers or designed with them in mind, females may experience frustration and anxiety and consequently be less satisfied with the website (Lin and Hsieh 2016). Additional research evaluating the influence of gender differences on website design and usability is needed to better inform website developers to be more female friendly.

The main goal of this evaluation was to identify which aspects of the website affected a user's ability to successfully navigate and use the website. Initial findings highlighted aspects of the PINEMAP DSS that served as

stumbling blocks in user navigation, and they also identified key differences between males and females in their eye movements and in their success in completing tasks. On the basis of the initial findings from phase 1, the PINEMAP DSS was modified to address the identified stumbling blocks. This evaluation aimed to determine whether the modifications made the PINEMAP DSS easier to use and whether the differences between males and females found in phase 1 persisted after implementing these changes. Specifically, the guiding research questions of this study were as follows:

1) Are there gender differences in successful navigation of the PINEMAP DSS?
2) Are there gender differences in successful interpretation of the information?
3) In what other ways are gender differences manifested?

2. Methods

a. Study design

The evaluation included two phases: phase 1, which was the exploratory, initial evaluation, and phase 2, which was the follow-up evaluation. Phase 1 involved experts from the forestry field because these respondents most closely represented the target audience of the PINEMAP DSS. Because of the difficult nature of recruiting an identical sample population, phase 2 involved an undergraduate student sample of convenience instead of forestry professionals.

All study participants from phase 1 and phase 2 were given 5 min to freely explore the PINEMAP DSS website. The free exploration period allowed participants to become familiar with the website and gave them an opportunity to read about the PINEMAP DSS or to begin using the tools. At the end of the allotted 5 min, participants were automatically advanced by the eye-tracking software to the second portion of the study, which measured the usability of the PINEMAP DSS by asking participants to complete one task and answer two multiple choice questions for each tool in the evaluation.

Three tools were evaluated: Extreme Minimum Temperature; Summer Precipitation; and Seedling Markets [later renamed to Cold-Tolerant Markets for Nurseries in phase 2 to better reflect the climate-based factors (winter minimum temperatures) used in the tool]. Altogether, there were three tasks and six multiple choice questions (see the appendix). The tasks were designed to utilize certain features of the PINEMAP DSS, one at a time, to measure the usability of individual website features. Additionally, each task built on the knowledge gained by successfully navigating the PINEMAP DSS to find and use

particular website features from the previous tasks. For example, the first task asked the users to input latitude and longitude coordinates, and the last task asked the users to enter the latitude and longitude coordinates, select a time period, and find a drop-down menu to turn on a specific map layer. Participants worked through this portion of the study at their own pace and advanced to the next task or question only when they were ready.

Participants' eye movements were recorded throughout their interactions with the PINEMAP DSS, from the free exploration period through the completion of all tasks and multiple choice questions, with an eye tracker. User demographics (age, gender, and education level) were also collected to determine if certain characteristics influenced a user's success in completing the tasks and/or answering the related questions correctly. In addition to the multiple choice questions testing user success with the activities, participants were asked open-ended questions about their experiences using climate data and climate information websites.

In addition to the study protocol already described, participants in phase 2 were asked about their experiences with activities linked to developing spatial skills and/or those that develop map reading skills such as scouting organizations, building blocks, playing video games, and using maps for navigation (Casey et al. 2008; Jirout and Newcombe 2015; Gold et al. 2018). They also completed a block-slicing test (Titus and Horsman 2009; Ormand et al. 2014) to measure one aspect of spatial ability, visual penetrative ability. These activities were added to the study design to help to determine whether differences in spatial ability and spatial skills could explain any potential differences between the eye movements and successful completion of tasks of males and females. All aspects of this study were carried out under approval through the Institutional Review Board for the Protection of Human Subjects.

b. Study populations

The participant group from phase 1 included 30 forestry students and professionals representing academia, industry, nonprofits, and federal and state governments (18 males and 12 females; 21–65 years old), and their highest degrees, either completed or in progress, ranged from undergraduate to doctoral degrees (6 undergraduate degrees, 17 master's degrees, and 6 doctoral degrees). They were recruited at a forestry conference in Durham, North Carolina, and were compensated afterward with their choice of a prepackaged candy bar or a rain gauge. The participant group from phase 2 included 12 undergraduate students (4 males and 8 females; 18–21 years old) in STEM and non-STEM majors, and their grade levels ranged from freshmen to juniors (2 freshmen,

8 sophomores, and 2 juniors). They were recruited from introductory physical science classes at a large research institution in the southeastern United States and were compensated afterward with a prepackaged candy bar. While the participants are not perfectly identical in background or position at the time of the study, they are similar within their respective phases (forestry professionals and undergraduate students, respectively), and there are no biases in majors, job titles, employers, and so on between the genders.

c. Validity and limitations

To ensure validity, the tasks and questions were written with the help of experts in the content area, and the complete study was reviewed by multiple graduate students. Additionally, the authors are aware of the following limitations: the two study populations were not identical; not all participants are forest managers or actively making management decisions, therefore the participants are not necessarily all from the same population as the intended PINEMAP DSS end users; changes were made to the PINEMAP DSS website, some based on the findings from phase 1 and others resulting from continued DSS tool development, between the two evaluation phases, so the website was not identical between phase 1 and phase 2; and, because of limited time with participants, only one instrument was used to measure spatial ability. Further, the study sample size of males and females was limited because of the emphasis on the eye-tracking applications (typical sample sizes are less than 30 subjects in eye-tracking research) and may or may not be generalizable to the targeted PINEMAP DSS audience or broader public audiences. Future work is suggested to address the limitations in this current study.

d. Eye tracking

The primary method of evaluation used is eye tracking. Eye tracking is a technology that has been used for more than 100 years (Jacob and Karn 2003) to study eye movements. In its current state, eye tracking is a noninvasive, harmless technology that allows researchers to determine where, when, and for how long a user is looking at specific elements on a computer screen. Eye tracking has been used across various applications, including website usability studies (Bojko 2006; Holmqvist et al. 2011; Drost et al. 2015; Loyola et al. 2015; Wilson et al. 2016).

The eye-tracking hardware used to evaluate the PINEMAP DSS utilized two different types of instruments. The phase-1 setup consisted of two Tobii Pro X2–60 (60 gaze points per second) eye trackers used with portable laptop computers (at a forestry conference with forestry experts). The phase-2 setup consisted of one Tobii TX-300 (300 gaze points per second) eye tracker used with a stationary desktop computer (on campus with undergraduate students). Before eye movements were recorded, each participant underwent a calibration process and was seated such that the eyes were ~65 cm from the eye tracker, all to ensure accurate and consistent eye-tracking data.

After the eye-tracking data were collected, the Tobii Studio software provided both qualitative and quantitative outputs for further analysis. The qualitative outputs included heat maps—products that indicate where and how many visual fixations occur. These products are suitable for visual analysis and general descriptions of user eye movements and behaviors but not for statistical analysis purposes. The quantitative outputs were calculated on the basis of areas of interest (AOIs), regions within the PINEMAP DSS that are defined by the researcher. For example, one AOI used in this study was the region in each of the PINEMAP DSS tools where the user entered in the latitude and longitude coordinates for a specific location. Three quantitative metrics were calculated specific to each AOI: the fixation count (FC), the total number of times that a user fixates anywhere within the AOI; the time to first fixation (TFF), the total number of seconds between the time at which the web page loads and the time at which the user fixates for the first time within the AOI; and the total visit duration (TVD), the total number of seconds over all visits to the AOI that the user fixates within the AOI. These metrics were exported from the Tobii software for statistical analyses. For quality control purposes, the eye-tracking data (i.e., FC, TFF, and TVD) from participants with low calibration rates ($<60\%$, with $N = 4$ from phase 1 and $N = 0$ from phase 2) were not included in the statistical analyses throughout this study.

e. Statistical analyses

The quantitative data—various eye-tracking metrics, scores from the multiple choice questions, and scores from the spatial ability measures—were examined for statistically significant differences across groups using many different tests. Most of the comparisons failed to show statistically significant differences across groups, likely because of the small sample sizes, especially when comparing across subgroups. For this reason, only effect sizes (Cohen's d) are presented to highlight large differences between groups where they exist.

3. Results and discussion

A major focus of the initial PINEMAP DSS eye-tracking evaluation and subsequent follow-up study was on identifying any differences in the overall usability of
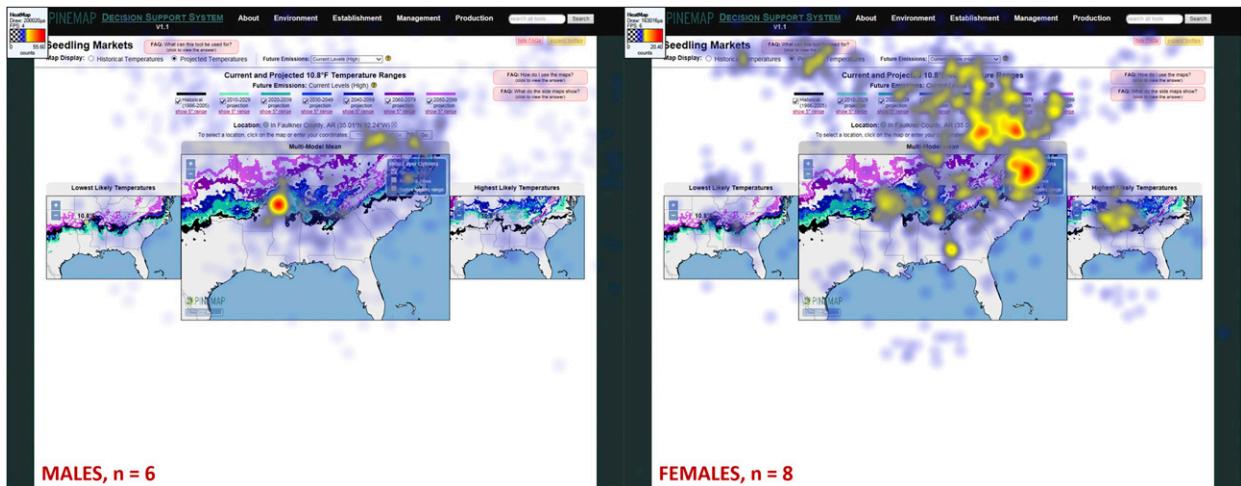
FIG. 1. Heat maps for a subset of the (left) males and (right) females for phase-1 Seedling Markets Tool (renamed to Cold-Tolerant Markets for Nurseries in phase 2). Red denotes high visual fixation counts, and blue denotes low visual fixation counts. Note that the maximum values for males and females are different. Males fixate more on the data within the center map, whereas females fixate more on the other website aspects.

the PINEMAP DSS between males and females. Eye tracking revealed qualitative differences in the visual fixation patterns of phase-1 males and females through-out many of the PINEMAP DSS features, but the most obvious differences were identified in the heat maps from the Seedling Markets tool (Fig. 1), the third of three tools used in the evaluation. For this particular tool, partici-pants were tasked with entering a specific latitude and longitude into the latitude and longitude box and toggling on various map layers by clicking the future emissions buttons located above the center map (Fig. 1). One of the questions associated with this task (see the appendix)

required that the participants look at the city located at the specified latitude and longitude and compare the information for the city among the three maps (the center map, showing the multimodel mean, and the two side maps, showing the lowest likely and highest likely temperatures on the left and right, respectively). The male visual fixation pattern showed a major hot spot (50+ visual fixations) over the city location in the center map and two smaller hot spots over the boxes where latitude and longitude coordinates were entered, indi-cating that the majority of male visual fixations were focused on the center map (Fig. 1). The female visual
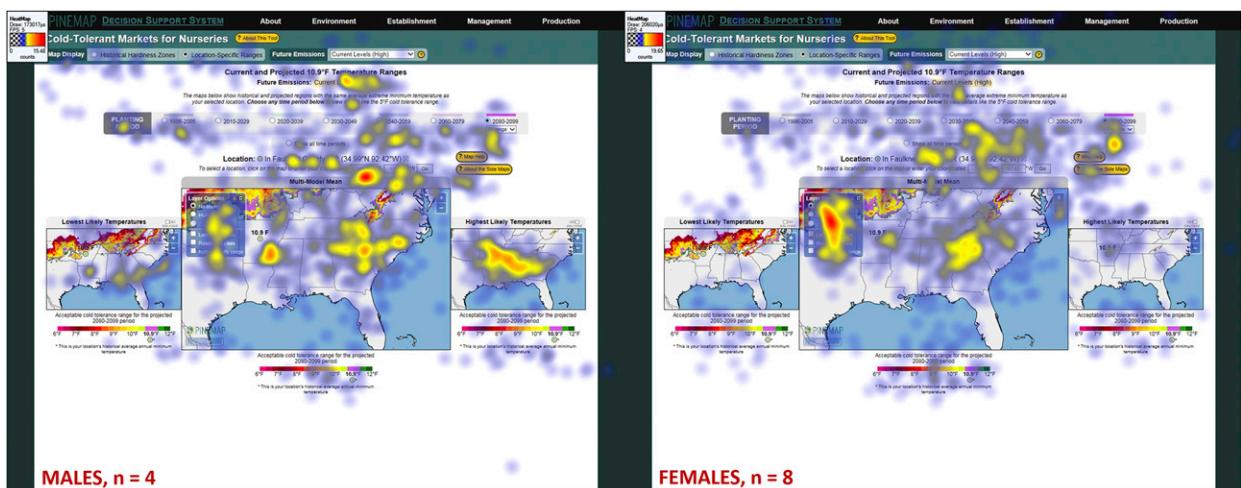

FIG. 2. Heat maps for (left) males and (right) females for the phase-2 Cold-Tolerant Markets for Nurseries (previously named Seedling Markets Tool in phase 1). Red denotes high visual fixation counts, and blue denotes low visual fixation counts. Males and females have similar visual fixation patterns for the content located above the center map, and both look at the data within the center map. Males give more attention to the map on the right, whereas females neglect the side maps almost entirely.

fixation pattern showed a broader range of hot-spot locations throughout the Seedling Markets tool (Fig. 1). Female visual fixations focused more heavily on the nonmap aspects of the website such as various features that could be toggled off and on, the latitude and longitude boxes, drop-down menus, and the map-layer options menu box. While some attention was given to the data, the majority was not.

In phase 2, the visual fixation patterns of males and females were more similar to each other than those from phase 1, but there were some key differences (Fig. 2). The phase-2 males and females had similar visual fixation patterns for the content located above the maps; however, females gave more attention to the center map and failed to give much attention to the side maps as compared with the males, meaning that females were not looking at all of the required data before answering the related question. In addition, the most attention given by females was to the map-layer options menu box (Fig. 2), which was left open by default in phase 2 (instead of minimized by default as it was in phase 1).

A notable difference between males and females in both phases is how they used their free exploration time at the start of the study. From the quantitative eye-tracking metrics, males fixated more on the data (the maps themselves) than did females, and in some cases, the females did not fixate anywhere within a given AOI (Fig. 3). In addition, females fixated more on the text on the home page as they read through the material than did males (e.g., 169 times vs 148 times, on average). During the tasks-and-related-questions portion of the study, males in both phases generally fixated more often than females in both phases on the map aspects that helped to answer the related questions at the end of the task (e.g., 14 times vs 6 times, on average, for the city location), and females generally fixated more often than males on other aspects of the DSS (e.g., 28 times vs 19 times, on average, for the map-layer options menu box) (Figs. 1 and 4). By giving more attention to the city location within the maps, males are likely receiving more information that will ultimately help them answer the related questions. By giving more attention to the website aspects above the maps, females are likely becoming more familiar with the various features of the specific tool and the website in general, but they are not gaining insight into the specifics of the data that will help answer the related questions. When combining these two major differences between males and females, how they use their free exploration time and where they focus their attention most during the tasks and while answering the related questions, differences
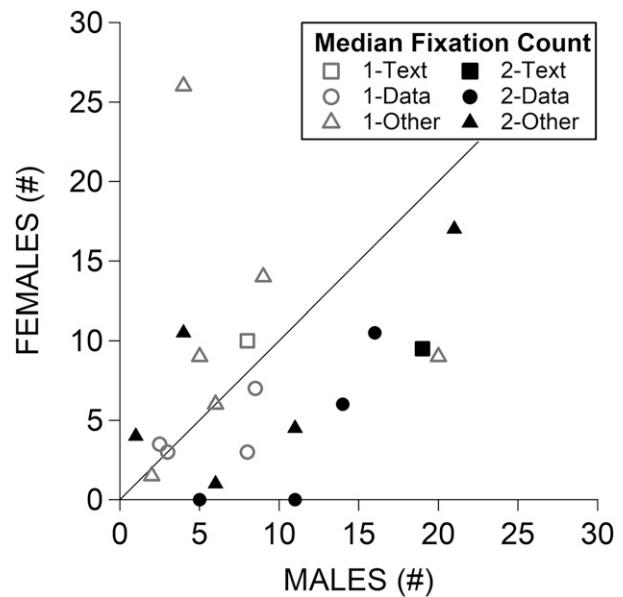

FIG. 3. A comparison of where males and females focus their attention during the 5 min of free exploration at the start of the study. Symbols represent the median value of visual fixation counts of males relative to females for individual areas of interest. Those above the diagonal line mean that females had a higher median visual fixation count than males, and those below the diagonal line mean that males had a higher median visual fixation count than females. Symbols are also differentiated between phase 1 (denoted with "1-" and open symbols) and phase 2 (denoted with "2-" and filled symbols) and by type of AOI, where "Text" refers to AOIs filled with blocks of text, "Data" refers to AOIs located within the maps, and "Other" refers to AOIs that include color bars, latitude and longitude boxes, and various other buttons.

in their success of answering the questions correctly can be expected.

Of the six questions asked in the tasks-and-related-questions portion of the study, males across both phases answered more questions correctly (Fig. 5), on average (males: mean = 4.74 and std dev = 0.96 vs females: mean = 4.21 and std dev = 1.32; Cohen's d = 0.40; small effect size). Of all subgroups, phase-2 males answered the most questions correctly, on average (mean = 5.25; std dev = 0.83), and phase-2 females answered the fewest questions correctly, on average (mean = 3.88; std dev = 1.69), resulting in a large effect-size difference between male and female performance in phase 2 (Cohen's d = 0.81). The three lowest performers from phase 2 (those with total scores less than or equal to 3 of 6) in this study were all females (Fig. 5). In addition, they also had the least experience with any activities related to developing spatial skills. Of the 12 spatial-related activities inquired about during phase 2 of this study, males had more experience and an average of 8.75 (std dev = 2.38),
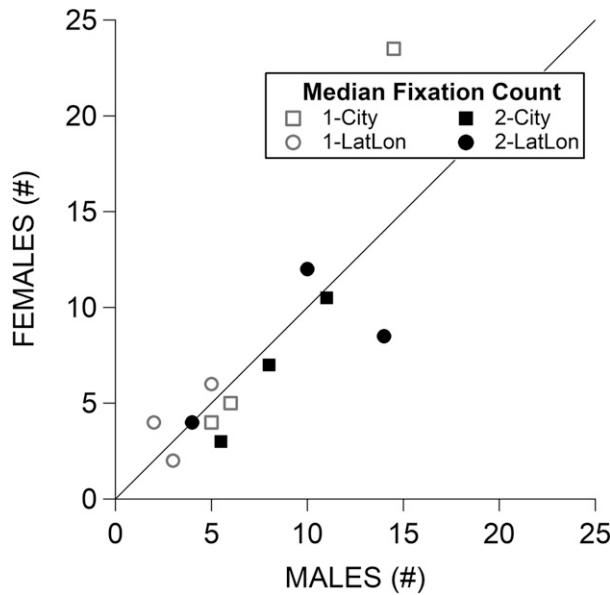
FIG. 4. As in Fig. 3, but showing a comparison of how and where males and females focus their attention during tasks and while answering the related questions. Here, "City" refers to the location of the specified city for each of the three tools, an indication that participants were looking at the map data, and "LatLon" refers to the latitude and longitude boxes above the maps, an AOI that is indicative of attention being given to the content above the maps.

females had less experience and an average of 6.38 (std dev = 3.16), and the three lowest performers (all female) had an average of 3.33 (std dev = 1.70). The lowest performers also had the lowest scores on the spatial ability test (mean = 3.33; std dev = 2.57) as compared with the average values for males (mean = 6.25; std dev = 5.80) and females (mean = 4.66; std dev = 3.82). Although experience with these types of activities and the spatial ability test used in this study do not directly

predict performance with the tasks-and-related-questions portion of the study, the similarities between the three lowest performers suggest these are factors that can influence performance.

4. Conclusions

The evaluation described here found differences in the overall usability of the PINEMAP DSS between males and females as measured by their performance on the tasks and multiple choice questions used in the evaluation. The eye-tracking data point to differences in where males look when interacting with various aspects of the PINEMAP DSS versus where females look, as well as the duration and frequency of their visual fixations on various PINEMAP DSS website elements. While not a direct indicator of usability and performance, low spatial ability and few experiences with activities linked to building spatial skills are associated with lower performance scores in this study.

The findings from this study highlight the influence end-user demographics and experience levels might have on the usability of DSS products and the need to consider these factors in the development of DSS products. For example, if DSS developers want to ensure females focus on the data found within the maps, extra effort should be given to ensure users are visually directed to that information (e.g., use an arrow or a pulsing banner to draw attention to the map). Similarly, if DSS developers want to ensure males read about a tool before using it, extra effort should be given to drawing their attention to the relevant text (e.g., include the necessary information on the same page as the tool itself and highlight it with a pulsing banner).
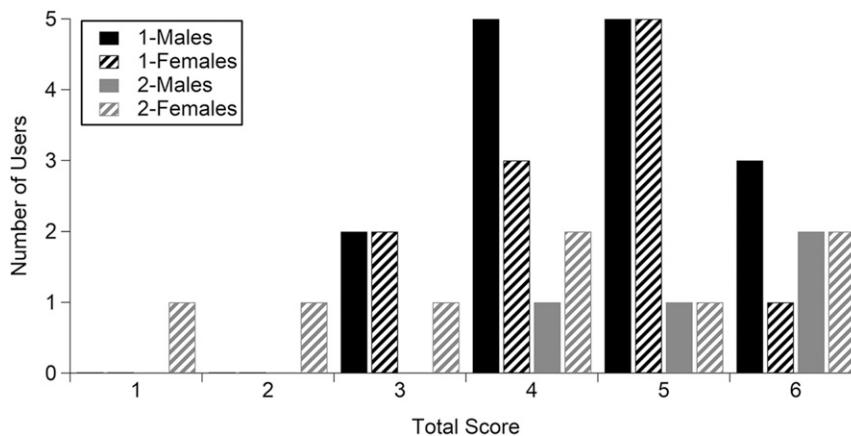

FIG. 5. Total score of a possible 6 for males and females in phase 1 (denoted with "1-") and phase 2 (denoted with "2-").

Additional research and effort are needed in the form of DSS evaluations to further address how gender might impact a user's success in navigating and interpreting information from a climate DSS, and other user demographics and experiences such as age and education level should also be considered. To ensure the development of a usable and easily understood climate DSS for decision-makers, the target audience, along with their demographics and experience levels, should be clearly identified, and evaluations that draw on existing evaluation frameworks (McNeal et al. 2014) should be included from the beginning of the design process (Perry et al. 2016).

APPENDIX

Tasks and Multiple Choice Questions

This appendix gives the tasks and multiple choice questions used in this study. It has not been edited to conform to American Meteorological Soci-ety journal style.

Task: Explore Historical Information

Navigate to the Extreme Minimum Temperature tool using the menu at the top of the page.
Once you're there, select the following options:
Temperature Threshold: 32°F
Map Display: Historical Average
Then answer the following question.
What is a typical value in central North Carolina?

- Between 65° and 75°F
- Between 65° and 75°C
- Between 65 days and 75 days
- Between 65 years and 75 years
- I am not sure.

Task: Explore Projected Change Information

Now select these options:
Map Display: Projected Change
Future Time Slice: 2020–2039
Future Emissions: Current Levels (High)
Location: Raleigh, NC (35.78°N, 78.64°W)
Then answer the following question.
What is the spread in the projected change of extreme minimum temperature occurrence?

- About 5 fewer days to about 15 fewer days

- About 5°F lower to about 15°F lower
- About 5 fewer days to about 25 fewer days
- About 5°F lower to about 25°F lower
- About 15 fewer days to about 25 fewer days
- About 15°F lower to about 25°F lower
- I am not sure.

Task: Explore Projected Average Information

Navigate to the Summer Precipitation tool and select the following options:
Map Display: Projected Average
Future Time Slice: 2080–2099
Future Emissions: Reduced Levels (Moderate)
Location: Columbia, SC (34.03°N, 80.90°W)
Then answer the following question.
What is the multimodel mean at this location?

- About 13°F
- About 13 inches
- About 15°F
- About 15 inches
- About 18°F
- About 18 inches
- I am not sure.

Task: Explore Time Series Plots

Now examine the time series plot at the bottom of the page.
Then answer the following question.
In the time series plot, which time slice has the greatest spread of possibilities?

- Historical
- 2020–2039 for Reduced (Moderate) Emissions
- 2040–2059 for Current (High) Emissions
- 2060–2079 for Reduced (Moderate) Emissions
- 2080–2099 for Current (High) Emissions
- I am not sure.

Task: Explore Projected Temperatures

Navigate to the Seedling Markets tool and select the following options:
Location: Little Rock, AR (34.74°N, 92.30°W)
Map Display: Projected Temperatures
Future Emissions: Current Levels (High)
Then answer the following question.
What is the trend with time from the Historical (black) to the 2080–2099 (magenta) time periods? How do these trends differ from the Lowest Likely to the Highest Likely Temperatures?

- Northward with more spread
- Southward with more spread

- Northward with less spread
- Southward with less spread
- I am not sure.

Task: Explore a Temperature Range

Next, show the 5° range for the 2080–2099 time period. Also, in the "Show Layer Options" menu, turn on the Native loblolly range layer.

Then answer the following question.

What is the mean projected location of the 5° range relative to the native loblolly range?

- It is mostly contained within the native loblolly range.
- It is mostly north of the native loblolly range.
- It is mostly south of the native loblolly range.
- I am not sure.

REFERENCES

Andreu, J., J. Capilla, and E. Sanchis, 1996: AQUATOOL, a generalized decision-support system for water-resources planning and operational management. *J. Hydrol.*, **177**, 269–291, https://doi.org/10.1016/0022-1694(95)02963-X.

Bojko, A., 2006: Using eye tracking to compare web page designs: A case study. *J. Usability Stud.*, **1**, 112–120.

Casey, B. M., N. Andrews, H. Schindler, J. E. Kersh, A. Samper, and J. Copley, 2008: The development of spatial skills through interventions involving block building activities. *Cogn. Instr.*, **26**, 269–309, https://doi.org/10.1080/07370000802177177.

Cyr, D., and C. Bonanni, 2005: Gender and website design in e-business. *Int. J. Electron. Bus.*, **3**, 565–582, https://doi.org/10.1504/IJEB.2005.008536.

Daron, J. D., S. Lorenz, P. Wolski, R. C. Blamey, and C. Jack, 2015: Interpreting climate data visualizations to inform adaptation decisions. *Climate Risk Manage.*, **10**, 17–26, https://doi.org/10.1016/j.crm.2015.06.007.

Djamasbi, S., T. Tullis, J. Hsu, E. Mazuera, and K. Osberg, 2007: Gender preferences in web design: Usability testing through eye tracking. *Proc. 13th Americas Conf. on Information Systems*, Keystone, CO, Association for Information Systems, 133, https://aisel.aisnet.org/amcis2007/133.

Drost, R., J. Trobec, C. Steffke, and J. Libarkin, 2015: Eye tracking: Evaluating the impact of gesturing during televised weather forecasts. *Bull. Amer. Meteor. Soc.*, **96**, 387–392, https://doi.org/10.1175/BAMS-D-13-00217.1.

Feuerriegel, S., and J. Gordon, 2018: Long-term stock index forecasting based on text mining of regulatory disclosures. *Decis. Support Syst.*, **112**, 88–97, https://doi.org/10.1016/j.dss.2018.06.008.

Gold, A. U., P. M. Pendergast, C. J. Orman, D. A. Budd, J. A. Stempien, K. J. Mueller, and K. A. Kravitz, 2018: Spatial skills in undergraduate students—Influence of gender, motivation, academic training, and childhood play. *Geosphere*, **14**, 668–683, https://doi.org/10.1130/GES01494.1.

Harold, J., I. Lorenzoni, T. F. Shipley, and K. R. Coventry, 2016: Cognitive and psychological science insights to improve climate change data visualization. *Nat. Climate Change*, **6**, 1080–1089, https://doi.org/10.1038/nclimate3162.

Holmqvist, K., M. Nyström, R. Andersson, R. Dewhurst, H. Jarodzka, and J. van de Weijer, 2011: *Eye Tracking: A Comprehensive Guide to Methods and Measures*. Oxford University Press, 560 pp.

Jacob, R. J. K., and K. S. Karn, 2003: Eye tracking in human–computer interaction and usability research: Ready to deliver the promises. *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research*, J. Hyönä, R. Radach, and H. Deubel, Eds., Elsevier, 573–605, https://doi.org/10.1016/B978-044451020-4/50031-1.

Jirout, J. J., and N. S. Newcombe, 2015: Building blocks for developing spatial skills: evidence from a large, representative U.S. sample. *Psychol. Sci.*, **26**, 302–310, https://doi.org/10.1177/0956797614563338.

Lin, C. J., and T.-L. Hsieh, 2016: Exploring the design criteria of website interfaces for gender. *Int. J. Ind. Ergon.*, **53**, 306–311, https://doi.org/10.1016/j.ergon.2016.02.002.

Loyola, P., G. Martinez, K. Munoz, J. D. Velasquez, J. Maldonado, and A. Couve, 2015: Combining eye tracking and pupillary dilation analysis to identify Website Key Objects. *Neurocomputing*, **168**, 179–189, https://doi.org/10.1016/j.neucom.2015.05.108.

McNeal, K. S., J. C. Libarkin, T. S. Ledley, E. Bardar, N. Haddad, K. Ellins, and S. Dutta, 2014: The role of research in online curriculum development: The case of *EarthLabs* climate change and earth system modules. *J. Geosci. Educ.*, **62**, 560–577, https://doi.org/10.5408/13-060.1.

Moss, G. A., and R. Gunn, 2007: Gender differences in website design: Implications for education. *J. Syst. Cybern. Inf.*, **5**, 38–43.

——, ——, and J. Heller, 2006: Some men like it black, some women like it pink: Consumer implications of differences in male and female website design. *J. Consum. Behav.*, **5**, 328–341, https://doi.org/10.1002/cb.184.

Ormand, C. J., C. Manduca, T. F. Shipley, B. Tikoff, C. L. Harwood, K. Atit, and A. P. Boone, 2014: Evaluating geoscience students' spatial thinking skills in a multi-institutional classroom study. *J. Geosci. Educ.*, **62**, 146–154, https://doi.org/10.5408/13-027.1.

Perry, S. C., and Coauthors, 2016: Get your science used—Six guidelines to improve your products. USGS Circular 1419, 46 pp., https://pubs.usgs.gov/circ/1419/c1419.pdf.

Pyke, C. R., B. G. Bierwagen, J. Furlow, J. Gamble, T. Johnson, S. Julius, and J. West, 2007: A decision inventory approach for improving decision support for climate change impact assessment and adaptation. *Environ. Sci. Policy*, **10**, 610–621, https://doi.org/10.1016/j.envsci.2007.05.001.

Roy, M., and M. T. H. Chi, 2003: Gender differences in patterns of searching the web. *J. Educ. Comput. Res.*, **29**, 335–348, https://doi.org/10.2190/7BR8-VXA0-07A7-8AVN.

——, R. Taylor, and M. T. H. Chi, 2003: Searching for information on-line and off-line: Gender differences among middle school students. *J. Educ. Comput. Res.*, **29**, 229–252, https://doi.org/10.2190/KCGA-3197-2V6U-WUTH.

Sahebjamnia, N., S. Ali Torabi, and S. Afshin Mansouri, 2017: A hybrid decision support system for managing humanitarian relief chains. *Decis. Support Syst.*, **95**, 12–26, https://doi.org/10.1016/j.dss.2016.11.006.

Salazar, K., 2010: Addressing the impacts of climate change on America's water, land, and other natural and cultural resources. Secretary of the Interior Order 3289, 4 pp., https://www.fws.gov/home/climatechange/pdf/SecOrder3289.pdf.

Simon, S. J., 2001: The impact of culture and gender on web sites: An empirical study. *Data Base Adv. Inf. Syst.*, **32** (1), 18–37, https://doi.org/10.1145/506740.506744.

Snover, A. K., N. J. Mantua, J. S. Littell, M. A. Alexander, M. M. McClure, and J. Nye, 2013: Choosing and using climate-change scenarios for ecological-impact assessments and conservation decisions. *Conserv. Biol.*, **27**, 1147–1157, https://doi.org/10.1111/cobi.12163.

Titus, S., and E. Horsman, 2009: Characterizing and improving spatial visualization skills. *J. Geosci. Educ.*, **57**, 242–254, https://doi.org/10.5408/1.3559671.

Wang, H.-F., 2014: Picture perfect: Girls' and boys' preferences towards visual complexity in children's websites. *Comput. Human Behav.*, **31**, 551–557, https://doi.org/10.1016/j.chb.2013.05.033.

Wilson, K. A., P. L. Heinselman, and Z. Kang, 2016: Exploring applications of eye tracking in operational meteorology research. *Bull. Amer. Meteor. Soc.*, **97**, 2019–2025, https://doi.org/10.1175/BAMS-D-15-00148.1.

Zhang, P., and N. Li, 2005: The importance of affective quality. *Commun. ACM*, **48** (9), 105–108, https://doi.org/10.1145/1081992.1081997.