Boise State University

## ScholarWorks

Geosciences Faculty Publications and Presentations

Department of Geosciences

6-2020

# Does Updating Natural Hazard Maps to Reflect Best Practices Increase Viewer Comprehension of Risk?
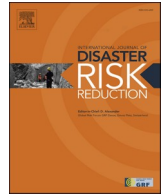
Carson C. MacPherson-Krutsky
*Boise State University*

Brittany D. Brand
*Boise State University*

Michael K. Lindell
*Boise State University*

# Does updating natural hazard maps to reflect best practices increase viewer comprehension of risk?☆

Carson C. MacPherson-Krutsky [*], Brittany D. Brand, Michael K. Lindell

*Department of Geosciences, Boise State University, USA*

A B S T R A C T

In this study, we examine whether updating an interactive hazard map using recommendations from the literature improves user map comprehension. Analyses of experimental data collected from 75 university students revealed that map comprehension scores were not significantly better for those who viewed a "best practices" map compared to those who viewed an existing version. This may be because the existing map was itself better than most other interactive maps. Additionally, we found map comprehension levels to have significant positive relationships with objective tests, but not self-reported measures of spatial ability. Moreover, self-reported spatial ability had statistically significant, but only moderately strong, correlations with objective tests. These results indicate that spatial ability should be measured objectively rather than through self-reported methods in research on map comprehension. Further research is needed to examine the cognitive processes involved in hazard map comprehension, especially using a broader range of map characteristics and population segments with more diverse cognitive abilities.

## 1. Introduction

Government agencies use hazard maps, in-print and online, to communicate environmental hazard risks. In many cases, maps made for use by experts such as geologists, engineers, land use planners, and emergency managers are shared with the public. However, these groups have diverse levels of hazard knowledge and cognitive abilities, which can produce confusion when maps contain technical or unnecessary information. As such, a one-size fits all approach to creating and disseminating maps for the purpose of communicating environmental hazards and risk is potentially problematic [1,2].

Despite their widespread use, few studies assess the usability of hazard maps, and even fewer studies have identified map characteristics that are essential for people to accurately assess their risks. Thus, research is needed to (1) determine how maps currently published on hazard management websites compare to the best available map display practices, as outlined in summaries such as Dodge et al. [3]; and (2) determine if people's map comprehension is a function of stable individual characteristics such as spatial, verbal, and numeric abilities.

Some progress toward addressing the issue of map usability can be drawn from the broader research literature on people's interpretations of maps—and even more broadly on visuospatial displays. However, most map studies examine people's map learning and memory and do not assess real-time inferences viewers draw from maps while they view them [4]. The lack of research on how people use and interpret hazard maps in real-time is an important limitation because that is typically how people use them.

The purpose of our study is to explore whether updating an interactive hazard map using best practices helps improve people's comprehension of risk. We also consider how individual differences in cognitive ability affect map comprehension. The results of our research inform strategies to better communicate environmental risks to the diverse audiences who can use map information to prepare for natural hazard events. With $2.6 billion spent annually on preparedness in the United States [5], it is imperative that maps used to communicate environmental hazard risks are effective.

## 2. Literature review

In the following section, we summarize research evaluating hazard maps, and then turn to a discussion of map types, cognitive processes in map comprehension, mapping best practices, and determinants of map

comprehension.

## 2.1. Hazard map studies

The Lindell [6] review of research on warnings of imminent hazards found a much more extensive literature on verbal elements of warnings than on graphic displays or numeric information. However, literature that assesses people's interpretations of hazard maps is increasing, especially for earthquakes [7], wildfires (e.g. Refs. [8–10]), volcanoes (e.g. Refs. [2,11,12]), floods (e.g. Refs. [13–17]), tornadoes [18–23], and hurricanes [24–33].

These publications explore a variety of dependent variables such as viewer perceptions of risk, risk area accuracy, preferences for map features, misconceptions about visualizations, and effects of user characteristics on performance. These studies concluded that risk area residents are better able to locate and orient themselves using aerial photographs and 3D maps with clearly labeled landmarks than with conventional contour maps [1,2,8,11] and that isarithmic maps produce better understanding than gradational shaded or binned maps. However, color coding scheme and probability coding (numerical vs. verbal) also influence participants' judgments, at least among geoscientists and emergency managers [12]. Furthermore, confusion can occur when aspects of the map are poorly defined, such as having too many or too few features, or have a confusing map legend [24,33]. In addition, people draw important inferences about risk information that is not explicitly provided [7].

Overall, the hazard map studies listed above signify the importance of assessing people's perceptions of map characteristics such as perceived relevance and ease of understanding, as well as accuracy of interpretation.

## 2.2. Map types

To better understand the broader literature, it is important to recognize that spatial displays, of which maps are a specific type, can be classified as iconic, relational, or hybrid [34]. An iconic display represents spatial objects. An example of an iconic display is a road map because it represents the network of roads and the locations of landmarks in a geographical area. A relational display, such as a graph, represents nonspatial variables such as average rainfall in each month of the year or the correlation between education and income. A hybrid display combines an iconic display (e.g., a base map) with a relational display to provide a spatial representation of nonspatial categories or quantities, as when temperature ranges are represented by map contours [35]. Thus, hazard maps are hybrid displays.

## 2.3. Cognitive processes in map comprehension

Accurate interpretation of a spatial display requires viewers to—(1) see the display clearly, (2) pay attention to relevant features, (3) develop a cognitive map, and (4) make inferences from their cognitive map to produce judgements, decisions, and actions [34]. The ability to see the display clearly is affected by factors such as visual element size and the degree of clutter in a display. Attention is influenced by "bottom-up" processes, in which visually salient features such as bright colors capture viewers' attention. It is also influenced by "top-down" processes in which viewers' expectations direct their attention to specific display elements. These expectations are generated by schemas, also known as mental models, which are generic belief structures about entities, their attributes, and the interrelationships among those attributes [36]. People can have schemas of varying comprehensiveness about maps in general and, in particular, about the specific map content being displayed. Accordingly, people can range in knowledge from novice to expert in each of these domains. Another important contributor to the encoding process is the viewer's spatial ability which, following Colom et al. [37]; can be defined as the ability to generate, retain, retrieve, and

transform visual images. Map inferences are determined by a viewer's goals, which can be self-generated (e.g., a desire to find the most direct route from one location to another) or externally imposed (e.g., an experimenter-assigned task to reproduce the map).

Most map research assesses the quality of the cognitive maps derived from physical maps or, to a lesser degree, from navigation through the environment. For example, many studies reviewed by Taylor [4] presented viewers with a map, asked them to study it, withdrew the map, and asked them to perform some task indicating the degree to which they learned the map's elements and their relationships (e.g., recall of landmarks, distances among points).

Only a few studies on map comprehension examine the basic elements of map reading skills [38–41]. Specifically, these are (1) *symbol recognition:* accurate interpretation of map symbols, (2) *direction finding:* the determination of geographical directions among landmarks using a map compass, and (3) *scale use:* determination of actual geographical distances among landmarks using a map scale. In addition, more sophisticated maps, such as topographical maps require (4) *contour utilization:* the determination of quantities such as elevations from the location of points within contours.

## 2.4. Mapping best practices

Maps can facilitate or impede viewers' map comprehension, depending upon the degree to which they are consistent with viewers' cognitive processes [11,12,42]. The impediments to map comprehension identified in the hazard map literature are consistent with a broader summary of the research literature on visual displays, which concludes that viewers' graph interpretations are a function of seven broad factors [43]. These factors include *data complexity* (e.g., the number of variables and categories within each variable), *data display characteristics* (e.g., the discriminability of graphical features—object positions, lengths/areas, colors, dimensionality), *viewer tasks* (e.g., retrieve point values, compare values, infer relationships), *viewer prior content knowledge* (expert vs. novice), *viewer prior knowledge of display conventions* (expert vs. novice), *visuospatial abilities*, and *working memory*.

### 2.4.1. Best practices for visual elements

Researchers have made a number of recommendations to increase map comprehension, such as best base map choice, most important map elements to display, appropriate symbols and labels, and clear hierarchical structure. For example, feature selection eliminates inessential map elements; visual salience draws viewers' eyes to the most important features [44,45]. There is also research that investigates the use of shape, size, and color of map symbols. In particular, shape ranges from abstract to iconic, with comprehension being fastest and most accurate for iconic symbols that do not need a legend [4]. Larger elements are easier to see and more readily attract attention, but can obscure other elements by cluttering the map if they are too large. Recommendations on color choice are outlined below.

Visual salience is often accomplished using color. There are five main recommendations for color choice. First, adapt color schemes to the type of data displayed, such as sequential schemes for data with increasing values (e.g. earthquake shaking intensities), diverging schemes for data whose values are above or below a critical value (e.g. temperatures above or below freezing), and qualitative schemes for nominal data (e.g. forest, lakes, and deserts are green, blue, and yellow, respectively) [12, 44,46]. Second, use seven or fewer color classes when displaying data because a greater number produces difficulty matching legend items with data layers [12]. Third, use color-blind friendly (CBF) colors schemes since 7–10% of the male population is red-green color-blind [8, 12,46]. Fourth, use real-life color to represent data when possible, such as blue for flooding and red for lava [15,44,45]. Finally, ensure that the colors in the legend match the colors on the map because transparency options and base map imagery can obscure or change map colors [45].

### 2.4.2. Best practices for content elements

Research on content choice has produced five recommendations. We use the term 'content' to refer to refer to verbal or numeric information provided on or next to a hazard map. First, content must be current and accurate [2]. If hyperlinks are broken, data are old, or information is no longer valid, map users may lose trust in the map and disregard the information—thus impeding personal preparedness [1,47]. Second, incorporate engaging auxiliary information to personalize the hazards [1,7,8,16]. Auxiliary information could include local photographs of past events, personal stories, infographics, and protection measures. Another way to personalize interactive maps specifically is to include a search by address function and the ability to zoom to locations of interest [1,9,13]. Third, avoid specialized terms that many people are likely to misunderstand, such as 100-year flood, peak ground acceleration, and debris flow [13,15]. Fourth, use easily understandable terminology to explain what each data layer and colored zone represents [45]. If this is done properly, users do not need to seek more information to understand the map. Fifth, avoid or clearly explain verbal labels for quantitative variables such as probabilities. Terms such as "low", "medium", or "high" are confusing because there is substantial variation in the numerical values that people assign to these labels [12,22]. This problem can be minimized by providing probabilistic information in multiple formats, supplementing verbal labels with probability percentages (e.g. 30% probability), natural frequencies (e.g. 3 in 10), or graphics such as risk ladders [48], pictographs [49], or shaded displays [12]. Since people vary in their ability to process probabilistic information, presenting more than one descriptor type allows a wider audience to understand the data.

### 2.4.3. An evaluation rubric for hazard maps

To develop the rubric, we conducted a literature review focused on effective map design, hazard maps as risk communication tools, and risk communication best practices. The review encompassed literature on both static and interactive maps, though most focused on static maps since fewer interactive map studies exist. The recommendations naturally separated into two categories, visual and content aspects of map design. Many of the recommendations were repeated in the literature so we consolidated them to create the "high performance" criteria of the evaluation rubric. We defined moderate and poor performance criteria from there.

The resulting rubric has two sections with nine visual and nine content elements. For each element a map can score from one (poor performance) to three (high performance) points. A map's total score is the points scored divided by the points possible. For example, a map that scores moderate on all items would have 36 points out of 54 possible for a total score of 0.67. The rubric can be used for multi-hazard or single-hazard maps and online or paper maps. Some rubric elements may not apply to every map. For example, visual rubric Element 6, "colors match hazard color," would not apply for an earthquake hazard map. In this case, the points for Element 6 would not be included in the total points possible. Table 1 summarizes the recommendations from the previous two sections for nine visual and nine content elements in the 'high performance' column of the hazard map evaluation rubric.

### 2.5. Determinants of map comprehension

#### 2.5.1. Cognitive abilities

Although some scholars suggest more complex models (e.g. Refs. [50,51]), propose that spatial abilities can be defined primarily by two factors, spatial visualization and spatial orientation. Spatial visualization is the ability to manipulate or transform the image of spatial patterns into other arrangements ([52]; p. 173). Spatial orientation is "the ability to perceive spatial patterns or to maintain orientation with respect to objects in space" ([52]; p. 149). In addition, a third spatial ability that seems particularly relevant to map comprehension is spatial scanning, which refers to "speed in exploring visually a wide or

**Table 1**

Visual and Content Elements for the 'high performance' Column of the Hazard Map Evaluation Rubric—see Appendix A for full rubric with references.

| Recommendations |
|---|
| **Visual** |
| V1. Aerial imagery base map used (or pops up as first map). |
| V2. Landmarks clearly visible to help viewers orient/locate themselves. |
| V3. Important map components are present and well-positioned on page. |
| V4. Visual hierarchy is achieved through appropriate colors, symbols, font size, line width, and other symbolization techniques. Most important map elements are emphasized. Base map is complementary, and does not distract from primary message. |
| V5. Appropriate color schemes used on all data—sequential for increasing values (intensities) diverging schemes for values above/below critical value (temperature - freezing), and qualitative for nominal data (trees, water, and desert are green, blue, and yellow, respectively). |
| V6. If applicable, map colors match hazard color. |
| V7. Fewer than 5 color classes used (7 or fewer is ideal). |
| V8. Legend colors are matched exactly with those on map. |
| V9. Color-blind friendly schemes are used. |
| **Content** |
| C1. Auxiliary information is present along with mapped data. |
| C2. Risk messaging is included and positively framed. |
| C3. Maps are personalized/customizable. |
| C4. Information appears to be accurate and up-to date and is presented in clear and concise manner. |
| C5. Protection measures are included along with risk. |
| C6. Jargon/specialized terms are not used in map or descriptions. |
| C7. Legend items are clearly explained. |
| C8. If data are probabilistic, both percent and natural frequencies are used and likelihood term is not used to describe the data. |
| C9. Qualitative (low-med-high) terms are not used. |

complicated spatial field" ([52]; p. 155).

Multiple studies find that individuals who have higher levels of spatial ability are better at interpreting and applying map information [38,53–55]. The types of spatial abilities that predict performance on spatial tasks depend on the scale of the representation. Specifically, spatial abilities at small (object) and large (environmental) scales are distinct even though they are positively correlated [56]. Environmental-scale tasks require a distinction between survey knowledge and route knowledge. Survey knowledge involves an allocentric perspective of map elements and their relationships (i.e., aerial view), whereas route knowledge involves an egocentric perspective (i.e., street view) that is defined by the sequence of steps required to move from one location to another [57]. Moreover, relevant spatial abilities also depend on the type of spatial task. For example, in studies of map utilization, the map is continuously present (e.g., Ref. [35]). By contrast, studies of map learning require the recall and reproduction of map elements (e.g., Refs. [57,58]).

Although there does not seem to be any research on this topic, it is also possible that map comprehension and spatial ability scores are affected by a user's level of verbal ability. Map comprehension tests and spatial ability tests require that test takers read or listen to verbal instructions about how to perform the task. As a result, complex instructions could depress scores on map comprehension or spatial tests for those with lower levels of verbal ability. If verbal ability is a significant predictor of map comprehension or spatial abilities, word choice becomes critical when designing experiments to test these factors.

Previous studies use a variety of instruments to measure cognitive abilities. These instruments separate into objective and self-reported abilities. Examples of objective cognitive tests include those developed by the Educational Testing Service [52] and Vandenberg and Kuse [59] that ask participants to perform various timed tasks. Each test measures a distinct cognitive ability. Instruments that measure self-reported or perceived abilities include the Santa Barbara Sense of Direction Scale (SBSOD, a measure of environmental-scale spatial ability), the Philadelphia Spatial Ability Scale (PSA, a measure of object-scale spatial ability), and the Philadelphia Verbal Ability Scale (PVA, a measure of

verbal ability) [53]. Since their development, both objective and self-reported styles of measurement have been used to investigate cognitive abilities [60–65]. Self-reported ability measures are much simpler to implement, but more research is needed to determine how well they correlate with objectively measured cognitive abilities.

### 2.5.2. Metacognition

One neglected research question is whether those who have greater levels of map comprehension are able to assess their performance and conclude that the task is easy, an assessment known as *metacognition* [66]. Although one might presume that metacognitive accuracy is a given—those who struggle to comprehend a map would be aware of the task's difficulty for them—this is not necessarily the case. There is ample support for precisely the opposite finding, the Dunning-Kruger effect, in which less competent people are oblivious to their own ignorance [67].

### 3. Research questions and hypotheses

The research reviewed in the previous sections leads to four research hypotheses (RHs) and two research question (RQs) that address the relationships of map comprehension, spatial abilities, and other cognitive abilities.

**RQ1**. Can map comprehension be meaningfully divided into a Basic Map Skill scale and an Advanced Map Skill scale?

**RH1**. Map comprehension scores of participants viewing a "best practices" hazard map will be significantly higher than those viewing an existing hazard map.

**RH2**. Objective spatial ability scores and self-report spatial ability scores will have significant positive correlations with each other but nonsignificant correlations with verbal ability.

**RH3**. SBSOD scores will have significant positive correlations with PSA scores but will have distinctly different correlations with other variables.

**RH4a-b**. Map comprehension scores will have significant positive correlations with (a) objective and (b) self-report spatial ability scores.

**RQ2**. Are map comprehension scores positively correlated with metacognitive awareness of performance?

### 4. Research design

### 4.1. Procedure

To test these research hypotheses and research questions, we randomly assigned participants to a two group between-subjects experimental design in which half of the participants viewed the conventional map and the other half viewed the best practice map (Fig. 1). Participants in both groups began by taking three timed objective tests of spatial abilities. After completing the spatial tests, participants logged on to the hazard map and answered a questionnaire. The questionnaire comprised a map comprehension quiz, three self-report spatial ability scales, a self-report verbal ability scale, and demographic questions. A total of 75 Boise State University students in introductory level courses participated in exchange for extra-credit toward their course grade. The protocol was approved by the Boise State University Institutional Review Board.

### 4.2. Hazard map development

Participants assigned to the existing map were directed to the *Oregon HazVu: Statewide Geohazards Viewer* (www.oregongeology.org/hazvu/), referred to below as Hazard Map 1 (HM1; Fig. 2). We selected this viewer as is currently in use, displays multiple hazards, and has procurable data layers.

We constructed the best practices hazard map by first developing a rubric consisting of best practices in hazard mapping and science communication from the literature described above (Table 1; see Appendix A for full rubric). We then applied the rubric to HM1 to identify areas of improvement that were then implemented to produce the "best practices" hazard map (HM2; bit.ly/dataview2) using ArcGIS Story Map software (Fig. 2). Finally, all hazard data in HM1 were imported to populate HM2. In addition to updating data colors and map legend terminology, HM2 also included a side-panel with auxiliary information, historical photos, definitions, and further explanation of legend items to help put the data in context. In all, HM2 involved 21 changes to HM1 (Appendix B). There were 15 specific changes in the visual criteria involving 7 of the 9 rubric items. In addition, there were 6 specific changes in the content criteria involving 6 of the 9 rubric items, with some addressing more than 1 rubric item.
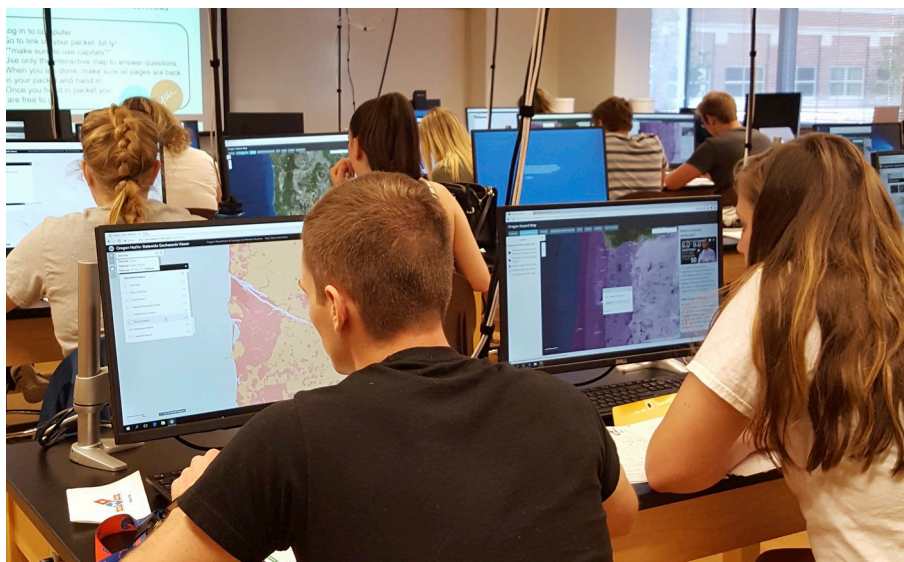


**Fig. 1.** After completing the timed spatial tests, students view HM1 (student on left) and HM2 (student on right) and fill out the map comprehension questionnaire.
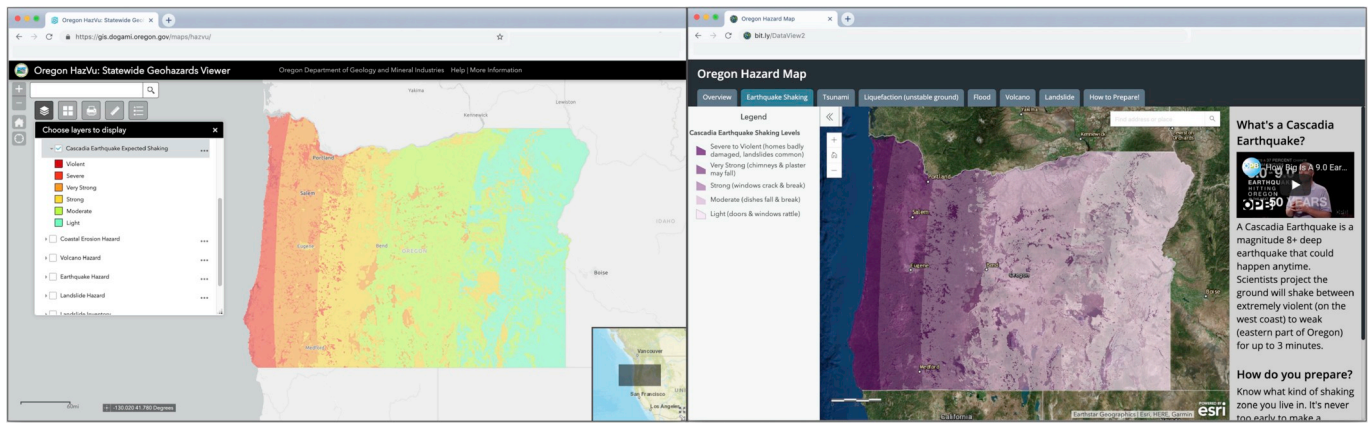
**Fig. 2.** Two screen captures show HM1 (left) and HM2 (right) with the earthquake shaking layer displayed. Note differences in color scheme, legend items, and auxiliary information.

### 4.3. Map comprehension, spatial ability, and cognitive ability measurement tools

The map comprehension scale comprised 13 questions in two categories covering the basic elements of map reading as well as more advanced skill in map interpretation (Table 2). Specifically, two items addressed participants' compass utilization, two items measured scale utilization, two items measured participants' ability to use the compass and scale in combination, two items measured legend utilization, and five items measured risk interpretation. The mean over the six items addressing compass utilization, scale utilization, and compass and scale in combination yielded a scale of Basic Map Skill. The mean over the seven items measuring legend utilization and risk interpretation yielded a scale of Advanced Map Skill. The internal consistency reliabilities for these two scales were $\alpha = .54$ and $.52$, respectively.

The three objective measures of spatial ability were selected from a series of cognitive tests published by Educational Testing Service—ETS [52]. The Paper Folding test measured visualization, the Cube

Comparison test measured spatial orientation, and the Map Planning test measured spatial scanning. The Paper Folding test requires people to select which of five options represents how a sheet of paper that has been folded and then hole-punched looks when it is unfolded. The Cube Comparison test requires people to determine if two cubes showing three faces with various designs, numbers, or letters visible on each face are different cubes or are the same cube that has been rotated to present different faces. The Map Planning test assesses people's ability to find the shortest route between two points in a stylized street grid that is partially obstructed by roadblocks. All three tests required the participants to answer as many questions as possible within 3 min and were hand-scored using the total number of correct responses for each test. The estimated reliabilities of these tests range 0.75–0.92 for Paper Folding, 0.77-0.89 for Cube Comparison, and 0.75-0.94 for Map Planning [52].

The three self-report spatial ability measures are the SBSOD and PSA [53,68], as well as the Allocentric View scale (Appendix C). The SBSOD and PSA scales contain questions describing the respondent's ability to perform a variety of tasks that require environmental- and object-scale spatial skills, respectively. For the SBSOD and PSA, participants responded to each item using a five-point Likert scales (*Strongly Agree* to *Strongly Disagree*) to indicate the degree to which it applied to them. These two spatial scales were supplemented by a newly developed Allocentric View scale that contains self-report items that are more directly related to map interpretation. That is, the items in this scale supplement the predominantly egocentric view items in the SBSOD. For the Allocentric View scale, participants responded to each item using a five-point scale (*Not at all to Very Great Extent*) to indicate its relevance to them.

Participants also completed the PVA self-report measure of verbal ability using a five-point Likert scales (*Strongly Agree to Strongly Disagree*) to indicate the degree to which each statement applied to them. Finally, they completed a Metacognition scale, which comprised a four items self-assessment of their performance on the map comprehension task. Participants used a five-point Likert scale (*Strongly Agree to Strongly Disagree*) to indicate the degree to which each statement applied to them.

After factor analysis and scale analysis, the SBSOD score was computed from the mean of all items except Item 9 ($\alpha = 0.89$), the PSA score was computed from the mean of Items 5–13 ($\alpha = 0.86$), the PVA score was computed using the mean of Items 1, 2, 6, and 7 ($\alpha = 0.64$), and the Allocentric View score was computed using the mean of all five items in that scale ($\alpha = 0.77$). The Metacognition score was computed using the mean of all four items in that scale ($\alpha = 0.77$). Variable labels are shown in Table 3.

**Table 2**
Map comprehension questions.

| Knowledge Category - *Question Focus* |
| --- |
| **Compass Utilization -** *Cardinal Directions: North, East, South, & West* |
| • Which of the following four cities is directly South of Portland? |
| • From Salem, OR, which direction would you have to travel to reach Dallas City, OR? |
| **Scale Utilization** - *Distances* |
| • How far is Salem (in Marion County) from Eugene (in Lane County) as the crow flies? |
| • Which two cities below are approximately 10 miles apart (as the crow flies)? |
| **Compass & Scale Utilization** - *Direction & Distance* |
| • If you travel about 10 miles East of Portland which town will you be in? |
| • Which direction and distance would you have to travel from Bend to Eugene as the crow flies? |
| **Legend Utilization** - *Hazard information shown in the legend & on the map* |
| • Which of the following cities could experience a tsunami? |
| • Eugene is expected to experience which level of shaking from a Cascadia earthquake? |
| • If you live at 701 Claggett St NE, Keizer, OR 97303 (at the corner of 7th Ave NE & Claggett St NE), which of the following hazards are likely to impact you? (choose as many as applicable) |
| **Risk Interpretation -** *Hazards & risk information associated with specific locations* |
| • If you are the owner of The Bank of America Financial Building (1001 SW 5th Ave, Portland, OR 97204 at the corner of 5th Ave. and SW Main St.), should you anticipate that <u>flooding</u> could impact your business in the next 100 years? |
| • If you are moving to Oregon and Mount Jefferson <u>volcano</u> is erupting, which of the following cities would be the most risky to live in? |
| • If your grandma lives at 3438 SE Chestnut St, Newport, OR 97366 (at the corner of SE 35th and SE Chestnut St.) and the <u>Cascadia earthquake</u> happens, which of the following is her home likely to experience? |
| • Rank locations Cloverdale, Pacific City, and Beaver from highest (Hi) to lowest (Lo) risk of being damaged from a <u>tsunami</u>. |

**Table 3**

Variable descriptions.

| Measure | Name | Type (range of values) |
|---|---|---|
| **Dependent Variables** | | |
| Basic Map Skill[a] | BasicSkill | Mean of values Q1-6 |
| Advanced Map Skill[a] | AdvancedSkill | Mean of values Q7-13 |
| **Independent Variables** | | |
| Map Type | MapType | HM1 (0), HM2 (1) |
| ETS Cube Comparison Test | CubeCompare | Score (max possible = 42) |
| ETS Map Planning Test | MapPlanning | Score (max possible = 40) |
| ETS Paper Folding Test | PaperFold | Score (max possible = 20) |
| SBSOD Scale | SBSOD | Mean of items except 9[a] (1–5) |
| PSA Scale | PSA | Mean of items 5–13[a] (1–5) |
| PVA Scale | PVA | Mean of items 1, 2, 6, & 7[a] (1–5) |
| Metacognition | Metacog | Mean of items[a] (1–5) |
| Allocentric View | AlloView | Mean of items[a] (1–5) |

[a] (high value = high perceived ability).

## 5. Results

### 5.1. Mean comparisons

The tests associated with RQ1—*Can map comprehension be meaningfully divided into a Basic Map Skill scale and an Advanced Map Skill scale?*—showed that scores on Basic Map Skill (Mean, $M = .81$) are significantly higher ($t_{71} = 2.14$, $p < .05$) than those on Advanced Map Skill ($M = 0.74$) and, as indicated in Table 4, the two scales have a significant Pearson correlation ($r = 0.23$) and a nonsignificant Spearman correlation with each other ($r = 0.20$). The small magnitude of both correlations suggests that map comprehension can be meaningfully divided into two relatively distinct skills.

### 5.2. Correlation analyses

To test the relationships between variables, we computed both Pearson and Spearman correlations (Table 4). We included Spearman correlations since the individual items cannot be assumed to be strictly interval or ratio level measures. However, discrepancies between statistically significant Pearson and Spearman Correlations are between 0.01 and 0.06. Upon testing the 95% confidence intervals for each discrepancy, we found these differences to be nonsignificant ($p > .05$). As such, the following results reference the Pearson correlation values.

Contrary to RH1—*Map comprehension scores of participants viewing a "best practices" hazard map will be significantly higher than those viewing an existing hazard map*—Table 4 shows that Map Type is significantly correlated only with Basic Skills and, unexpectedly, that correlation is negative ($r = -.27$). That is, participants who viewed HM2 tended to have lower Basic Map Skill scores than those who viewed HM1. Moreover, Map Type also has significant negative correlations with ETS Map Planning ($r = -.24$) and Allocentric View ($r = -0.24$).

Mostly consistent with RH2—*Objective spatial ability scores and self-report spatial ability scores will have significant positive correlations with each other but nonsignificant correlations with verbal ability*—the three ETS spatial ability tests have significant positive correlations with each other (average correlation, $\bar{r} = 0.49$) and all three have significant positive correlations with PSA ($\bar{r} = 0.35$), and Allocentric View ($\bar{r} = 0.31$). However, Map Planning has the highest correlations with these two variables ($r = .41$ and .36, respectively) and also with SBSOD ($r = 0.26$). Neither Cube Comparisons nor Paper Folding is significantly correlated with SBSOD. Although not hypothesized, the ETS Cube Comparisons and Map Planning tests have significant positive correlations with Metacognition ($\bar{r} = .32$). Contrary to the hypothesis, PVA score has significant positive correlation with PSA ($r = 0.26$).

Partially consistent with RH3—*SBSOD scores will have significant positive correlations with PSA scores but will have distinctly different correlations with other variables*—SBSOD and PSA have a significant positive correlation ($r = 0.52$). Unexpectedly, however, they have similar positive correlations with Allocentric View ($\bar{r} = 0.56$) and Metacognition ($\bar{r} = 0.35$) The only notable difference in their patterns of correlations is that PSA is more strongly correlated with PVA ($r = 0.25$ vs. 0.07), but neither of these correlations is statistically significant.

Partially consistent with RH4—*Map comprehension scores will have significant positive correlations with objective spatial ability scores*—both Basic Map Skill ($r = .28$) and Advanced Map Skill ($r = 0.32$) have significant positive correlations with Map Planning. However, only Advanced Map Skill has a significant positive correlation with Paper Folding ($r = .29$). and neither map comprehension scale has a significant correlation with Cube Comparison.

Contrary to RH5—*Map comprehension scores will have significant positive correlations with self-report spatial ability scores*—the correlations of both map comprehension scales with all self-report spatial ability scales are nonsignificant.

The tests associated with RQ2—*Are map comprehension scores positively correlated with metacognitive awareness of performance?*—show that Metacognition has a significant positive correlation with Advanced Skills ($r = .26$) but not Basic Skill ($r = 0.21$), although the difference between these two correlations is not statistically significant. Although not hypothesized, Metacognition and Allocentric View have significant positive correlations with each other ($r = 0.37$).

### 5.3. Ordinary least squares (OLS) regression analyses

To further test the results from RH1 and RH4, Map Type and Map Planning were entered as potential predictors of Basic Map Skill and Advanced Map Skill. Table 5a shows the results of the analyses for the prediction of Basic Map Skill. The left-hand panel of table shows that, after entering Map Type at the first step, Map Planning failed to enter after that. Conversely, the right-hand panel shows that, after entering

**Table 4**

Mean (*M*), standard deviations (*SD*), & Pearson (lower left) and Spearman (upper right) correlations ($r_{ij}$) among variables.

| | M | SD | Map Comp | | | ETS tests | | | Self-report scales | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | 1. Basic | 2. Adv. | 3. Type | 4. Cube | 5. Map | 6. Paper | 7. SBSOD | 8. PSA | 9. PVA | 10. Meta | 11. Allo |
| 1. BasicSkill | .81 | .21 | – | .20 | -.28* | .15 | .27* | .09 | .04 | .09 | .02 | .13 | .21 |
| 2. AdvancedSkill | .74 | .29 | .23* | – | .01 | .13 | .26* | .34** | -.04 | -.05 | .24* | .28* | .20 |
| 3. MapType | .49 | .50 | -.27* | .00 | – | -.16 | -.26* | .08 | -.09 | -.16 | -.03 | -.17 | -.25* |
| 4. CubeCompare | 24.04 | 8.06 | .17 | .17 | -.17 | – | .58** | .46** | .25* | .24* | -.19 | .27* | .25* |
| 5. MapPlanning | 22.59 | 8.02 | .28* | .32** | -.24* | .55** | – | .49** | .28* | .43** | .11 | .37** | .37** |
| 6. PaperFold | 12.35 | 3.69 | .08 | .29* | .07 | .48** | .44** | – | .20 | .43** | .16 | .22 | .32** |
| 7. SBSOD | 3.20 | .89 | .07 | -.08 | -.11 | .21 | .26* | .19 | – | .46** | -.02 | .33** | .49** |
| 8. PSA | 3.18 | .91 | .08 | -.09 | -.17 | .25* | .41** | .38** | .52** | – | .23 | .39** | .55** |
| 9. PVA | 2.90 | .87 | .01 | .22 | -.04 | -.15 | .19 | .15 | .07 | .26* | – | -.02 | .11 |
| 10. Metacog | 3.61 | .98 | .21 | .26* | -.18 | .26* | .38** | .18 | .32** | .38** | .01 | – | .37** |
| 11. AlloView | 2.80 | 1.03 | .21 | .23 | -.24* | .26* | .36** | .31** | .53** | .58** | .17 | .37** | – |

*$p < .05$, **$p < .01$ (2-tailed).

**Table 5**

Regression of Basic and Advanced Map Skill Scores onto Map Type and Map Planning Scores.

| | Map Type Entered First | | | | | Map Planning Entered First | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | b* | SE(b) | β | t | Sig. | b* | SE(b) | β | t | Sig. |
| **a. Regression of Basic Map Skill Scores** | | | | | | | | | | |
| Constant | .87 | .033 | | 26.48 | .00 | .64 | .074 | | 8.59 | .00 |
| Map Type | -.11 | .047 | -.27 | −2.40 | .02 | | | | | |
| Map Planning | | | | | | .01 | .003 | .28 | 2.48 | .02 |
| | Adj $R^2$ = .06 | | | | | Adj $R^2$ = .07 | | | | |
| | $F_{1,73}$ = 5.78 | | | | | $F_{1,73}$ = 6.15 | | | | |
| | $p$ = .02 | | | | | $p$ = .02 | | | | |
| **b. Regression of Advanced Map Skill Scores** | | | | | | | | | | |
| Constant | .44 | .111 | | 3.94 | .00 | .47 | .097 | | 4.89 | .00 |
| Map Type | .04 | .067 | .08 | 0.66 | .51 | | | | | |
| Map Planning | .01 | .004 | .34 | 2.94 | .01 | .01 | .004 | .32 | 2.87 | .01 |
| | Adj $R^2$ = .08 | | | | | Adj $R^2$ = .09 | | | | |
| | $F_{1,73}$ = 8.62 | | | | | $F_{1,73}$ = 8.24 | | | | |
| | $p$ = .01 | | | | | $p$ = .01 | | | | |

Note. b* denotes the unstandardized regression coefficient; *SE(b)* denotes the standard error of the regression coefficient; β denotes standardized regression coefficient.

Map Planning at the first step, Map Type, failed to enter after that. Table 5b shows the results of the analyses for the prediction of Advanced Map Skill. Map Type failed to enter at the first step (not shown) but, as shown in the left panel, Map Planning did enter in the second step while Map Type remained nonsignificant. Conversely, the right-hand panel shows that, after entering Map Planning at the first step, Map Type, failed to enter after that.

To further test RH2, the self-report measures were entered as potential predictors of ETS scores. Table 6 shows that only PSA scores significantly predicted Paper Folding test scores (Adj $R^2$ = 0.10 in the left-hand panel) and Map Planning test scores (Adj $R^2$ = 0.14 in the right-hand panel), but not Cube Comparison scores (Adj $R^2$ = 0.04 in the center panel). SBSOD scores did not significantly predict any of the ETS scores.

The validity of OLS regression analyses depends upon four assumptions—(1) linearity of the relationships between the independent and dependent variables, (2) independence of errors, (3) homoscedasticity (constant error variance), and (4) normal distribution of errors. Tests following the procedures in Ott and Longnecker [69]; Chapter 13) were conducted for the data used in the regression analyses above and revealed that Assumption 1 is supported by scatterplots of map comprehension against each of the independent variables, which revealed no indication of curvilinearity. Moreover, Assumption 2 is reasonable because the data are cross-sectional so there is no serial autocorrelation. Finally, Assumption 3 is supported by residual plots showing approximately constant dispersion across all values of the independent variables, and Assumption 4 is supported by linearity in the p-p plots of the standardized residuals.

## 6. Discussion

### 6.1. RQ1: can map comprehension Be meaningfully divided into a basic map skill scale and an advanced map skill scale?

The ability to interpret a hazard map is an important skill because many people need these hybrid visuospatial displays to determine whether they are in a hazard zone and, thus, need to take action to protect themselves from hazard impact. Basic and advanced map skills both require a degree of knowledge of mapping conventions and visuospatial skills. However, the results from the analyses of RQ1 suggest that these two types of map skills are somewhat distinct because there were significantly higher scores on basic skill than on advanced skill and the two scales were not significantly correlated. More generally, the fact that scores on Basic Map Skill (M = .81) were substantially less than perfect poses a challenge for developers of hazard maps because it means that people make errors when using the two most fundamental elements of these displays—the compass and scale. Further research is needed to determine if this lack of basic map skill can be replicated in samples that are more representative of the broader population. However, it seems likely that map comprehension scores will be even lower in a general population sample than in a university student sample that has been selected specifically for its higher level of cognitive ability. If so, research will also be needed to identify the specific impediments to successful compass and scale utilization, and either develop training methods to improve basic skill or create displays that overcome these impediments.

### 6.2. RH1: map comprehension scores of participants viewing a "best practices" hazard map will Be significantly higher than those viewing an existing hazard map

The lack of support for RH1 is quite surprising because Map Type not only had nonsignificant correlation and regression coefficients with

**Table 6**

Regression of ETS scores onto self-report spatial scale scores.

| | Paper Folding | | | | | Cube Comparison | | | | | Map Planning | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | b* | SE(b) | β | t | Sig. | b* | SE(b) | β | t | Sig. | b* | SE(b) | β | t | Sig. |
| Const | 7.52 | 1.73 | | 4.35 | .00 | 15.58 | 3.91 | | 3.99 | .00 | 10.18 | 3.69 | | 2.76 | .01 |
| SBSOD | .06 | .53 | .02 | .12 | .90 | 1.15 | 1.20 | .13 | .96 | .34 | .67 | 1.13 | .07 | .59 | .56 |
| PSA | 1.45 | .53 | .35 | 2.72 | .01 | 1.50 | 1.21 | .17 | 1.24 | .22 | 3.23 | 1.14 | .36 | 2.83 | .01 |
| | Adj $R^2$ = .10 | | | | | Adj $R^2$ = .04 | | | | | Adj $R^2$ = .14 | | | | |
| | $F_{2,72}$ = 5.25 | | | | | $F_{2,72}$ = 2.49 | | | | | $F_{2,72}$ = 6.82 | | | | |
| | $p$ = .01 | | | | | $p$ = .09 | | | | | $p$ = .002 | | | | |

Note. b * denotes the unstandardized regression coefficient; *SE(b)* denotes the standard error of the regression coefficient; β denotes standardized regression coefficient.

Advanced Map Skill, it had a significant *negative* correlation with Basic Map Skill. A possible methodological explanation for the nonsignificant correlation and regression coefficients with Advanced Map Skill is that this variable has only modest reliability ($\alpha$ = .52), which would attenuate its correlation with other variables [70]. However, this explanation is contradicted by the finding that Advanced Map Skill had significant correlations with other variables, so this scale seems to be measuring a meaningful construct even though its reliability is lower than is desirable. In any event, the map comprehension scales need further development to increase their psychometric quality.

An alternative explanation for the nonsignificant difference between map types is that there was essentially no meaningful difference between the two map types with respect to their demands for Advanced Map Skill. One variation of this explanation is that the changes made in transforming HM1 to HM2 were an inadequate operationalization of "best practices". Although this possibility cannot be ruled out definitively, it seems unlikely because—as noted above—the production of HM2 involved an extensive set of changes. A second variation of this explanation is that HM1, the existing map, was already quite good at meeting the participants' information needs with respect to advanced map skill, so the improvements implemented in HM2 had a minimal psychological impact on the participants. This explanation is consistent with the finding that HM1 already met many of the best practices. Thus, to better address this issue, further research should examine people's ability to process the information from hazard maps that encompass a wider range of quality with respect to the rubric elements in Table 1.

The explanation for the negative correlation of Map Type with Basic Map Skill involves the software used to create HM2, which was based on uploaded and formatted data and content in ArcGIS Story Maps. Story Maps software has many options but also has feature display limitations. For example, this software sets the map legend to pop-up *only when clicked.* As the first author watched people navigate HM2, it was apparent that many of them failed to click on the legend, which makes accurate interpretation almost impossible. By contrast, HM1 had a legend always visible. In addition, Story Maps also makes the scale bar a specific color independent of the base-map. Consistent with recommendations from previous studies, HM2 included an aerial image base map and the scale bar was dark grey. This made seeing the scale bar a bit challenging. By contrast, HM1 had a more visible scale bar and included measurement tools that could be used to measure distances precisely. Since the map comprehension test included questions about distance, this would also have contributed to slightly higher scores for HM1 viewers on Basic Map Skill.

### 6.3. RH2: objective spatial ability scores and self-report spatial ability scores will have significant positive correlations with each other but nonsignificant correlations with verbal ability

The partial support for RH2 is consistent with previous research. Specifically, the PSA has moderately high correlations with Map Planning ($r$ = 0.39) and Cube Comparisons ($r$ = 0.36), and a noticeably lower, but still significant, correlation with Paper Folding ($r$ = 0.23). By contrast the SBSOD had noticeably lower correlations with the three ETS tests ($r$ = 0.26, 0.19, and 0.21, respectively). These results support the contention that the SBSOD and PSA, though highly correlated ($r$ = 0.51), are indeed measuring somewhat different constructs [56].

Moreover, consistent with RH2, there are nonsignificant correlations of Paper Folding ($r$ = −0.04), Cube Comparisons ($r$ = 0.14), Map Planning ($r$ = −0.14), and SBSOD ($r$ = 0.06) with PVA. However, contrary to this hypothesis, PVA has a significant positive correlation with PSA ($r$ = 0.25). It is not obvious why this is the case because all three of the ETS spatial ability tests and the SBSOD have instructions that are at about the same level of verbal complexity as those for the PSA. Thus, further research is needed to determine if this finding can be replicated and, if so, explained.

As a practical matter, the poor predictability of the ETS tests from the

SBSOD and PSA, as shown in Table 6, is unfortunate because the ETS tests are timed and, therefore, must be administered in a carefully controlled setting such as a laboratory. By contrast, the SBSOD and PSA are untimed and can be administered in an uncontrolled setting such as a mail or Internet survey. In turn, this restriction in ETS test administration limits the types of population segments that can be tested using these scales. Consequently, further studies of the effects of spatial abilities on map comprehension should administer the ETS tests in controlled settings.

### 6.4. RH3: SBSOD scores will have significant positive correlations with PSA scores but will have distinctly different correlations with other variables

Regarding RH3, the high correlation of the SBSOD and PSA is consistent with the Hegarty et al. [56] conclusion that these two scales measure related but distinct types of spatial ability—the SBSOD measures spatial ability at the environmental scale (e.g., wayfinding) and the PSA measures spatial ability at the object scale (e.g., object manipulation). The support for this conclusion is particularly noticeable in the factor loadings in Appendix C. Moreover, the only significant correlation of the SBSOD with an ETS test is with the Map Planning test—the only one of these tests that assesses a skill approximating wayfinding at the object scale. Nonetheless, it is difficult to explain, given the assumption that the PSA measures object-scale spatial ability, that this scale's highest correlation with an ETS test is also with the Map Planning test. The most logical explanation is that performance on the Map Planning test draws upon spatial ability at both the object and environmental scales. The present study extends this finding by showing that the PSA scale and Map Planning test have similar patterns of correlations with Allocentric View, and Metacognition, all of which have significant positive correlations with each other. However, the present results provide no support for the contention that the SBSOD and PSA have distinctly different correlations with other variables.

### 6.5. RH4a-b: map comprehension scores will have significant positive correlations with (a) objective and (b) self-report spatial abilities scores

Partially consistent with RH4a, Map Planning was significantly correlated with Basic Map Skill ($r$ = .27) and Advanced Map Skill ($r$ = 0.32). In addition, Paper Folding was significantly correlated with Advanced Map Skill. ($r$ = 0.29) but not Basic Map Skill ($r$ = 0.08). However, Cube Comparison was not significantly correlated with either measure of map comprehension. These results suggest that the Map Planning test provides the most direct measure of the cognitive skills required for map comprehension.

Contrary to RH4b, neither the SBSOD nor the PSA was significantly correlated with Basic Map Skill ($r$ = 0.07 and 0.08, respectively) or Advanced Map Skill ($r$ = −0.08 and −0.08, respectively). Indeed, even the Allocentric View scale, which was constructed to be a self-report scale of map comprehension, lacked statistically significant correlations with the two map comprehension measures. The Allocentric View scale does not appear to have suffered from variance restriction ($SD$ = 1.03 is approximately 20% of the scale range) or attenuation due to unreliability ($\alpha$ = 0.77), but there is some room for improvement in this scale and, as noted earlier, substantial room for improvement in the psychometric quality of the map comprehension scales.

### 6.6. RQ2: do participants have a metacognitive awareness of their performance on map skills?

The results regarding metacognitive awareness showed that participants' assessments of their performance is significantly correlated with Advanced Map Skill. That is, those who were better at this task were able to assess their performance and conclude that the task was easier. This metacognitive accuracy is the opposite of the Dunning-Kruger effect, in

which less competent people are oblivious to their own ignorance [67]. This finding suggests feedback from the task itself provided poor performers with an assessment of the quality of their performance. In turn, this suggests that map users who are experiencing difficulty are likely to recognize their need to use general Help tabs if these are readily accessible. Indeed, the lower performance associated with the absence of a continuously visible map legend in HM2 suggests that context-dependent help features would be a particularly useful addition to hazard maps.

*6.7. Study limitations & opportunities for future work*

The first study limitation is the sample; students are a subset of the general population that can be assumed to have higher levels of verbal and numeric abilities because they are explicitly selected for admission on the basis of these cognitive abilities. However, it is less clear whether they have higher levels of spatial ability because universities do not use this cognitive ability as an explicit selection criterion. If university students do indeed have generally higher levels of spatial ability, then the absence of those who score low on this ability would produce a reduced variance and, in turn, attenuate the estimates of the correlation in the general population [70]. Thus, it is possible that use of a student sample underestimates the magnitude of the correlations found in this study. To overcome this sampling bias, future map comprehension studies should aim to recruit participants with a broader range of ages and abilities to be more representative of the population using these maps. With a more representative sample, we would expect larger correlations between variables. In practice, people may view hazard maps with a family member or friend, so future research could also include testing map comprehension in pairs or groups. Group discussion has been shown to improve reading comprehension [71] and may also improve map comprehension.

A second issue associated with this sample is that the students were not residents of the mapped area. This lack of familiarity with the area might have depressed map comprehension scores, especially for those with low spatial ability. To address this issue, future research on map comprehension should be conducted using samples of people who live in the mapped area.

A second study limitation arises from the type of map studied. Specifically, interactive hazard maps are fairly new, so this study is one of few investigating how people view and interpret dynamic map information. One consequence of the scarcity of prior studies on dynamic maps is that many of the recommendations used to update HM1 were made primarily for plan-form maps. It may be that people interpret maps differently when they are online versus in-print and that recommendations for one type do not apply well to the other. Thus, one future research objective should be to determine if providing the same hazard information on a plan-form and interactive map leads to comparable user comprehension levels.

A related issue is that, with the increasing use of interactive hazard maps, more research is needed on both single- and multi-hazard maps. Better understanding of how people navigate and use map features, how long they spend on the maps, and what kind of information they absorb are topics on which more research is needed. Assessment of the cognitive processes and cognitive abilities involved in map comprehension could also be expanded. More studies are needed to further identify which abilities predict map comprehension and how they are recruited in processing hazard maps [72].

The third limitation concerns whether the regression models are specified correctly. The available literature on map comprehension indicates that many, if not most, of the relevant variables have been included in the model, but the models in Tables 5 and 6 only account for ~4–14% of the variation in the dependent variables. This means either that the variables included need to be measured more reliably or that there are omitted variables that were not included in the analysis. The estimated reliabilities for SBSOD ($\alpha = 0.89$) and PSA ($\alpha = 0.86$) are quite satisfactory, but those for Basic Map Skill ($\alpha = 0.54$), Advanced Map Skill ($\alpha = 0.52$), and PVA ($\alpha = 0.64$) have ample room for improvement. With regard to omitted variables, it is possible that adding measures of numeric ability would improve the prediction of map comprehension. Further study is needed to test these variables and to identify additional predictors of map comprehension.

## 7. Conclusions

This study provides a practical test of whether hazard map design and content recommendations are necessary to improve user comprehension of risk. We found that a "best practices" interactive map provided no improvement over an original interactive map. This may be because the original interactive map scored higher on the rubric than many other interactive maps. Consequently, although HM1 might be as effective as the "best practices" map (HM2), other hazard maps may need to be improved to reach the same degree of comprehension. Thus, government agencies should design their interactive hazard maps for the public by addressing the rubric elements in Table 1.

As expected, objectively measured spatial ability is an important determinant of peoples' ability to interpret map information. Specifically, spatial scanning, as measured by the ETS Map Planning test, was a somewhat better predictor of both measures of map comprehension than was spatial orientation (Paper Folding) or spatial visualization (Cube Comparison). Unexpectedly, however, self-reported spatial ability does not significantly predict map comprehension and poorly predicts objectively measured spatial ability.

Many of the studies referenced above use individual perceptions of map objects and information to develop map recommendations. Our results suggest that more quantitative metrics may be better. Nonetheless, the regression analyses accounted for only a small portion of the variation in map comprehension. More research is needed to better assess the degree to which different factors contribute to high map comprehension levels.

## Acknowledgements

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijdrr.2020.101487.

## References

[1] D. Dransch, H. Rotzoll, K. Poser, The contribution of maps to the challenges of risk communication to the public, Int. J. Digital Earth 3 (3) (2010) 292–311, https://doi.org/10.1080/17538941003774668.

[2] R. Nave, R. Isaia, G. Vilardo, J. Barclay, Re-assessing volcanic hazard maps for improving volcanic risk communication: application to Stromboli Island , Italy, J. Maps 6 (1) (2010) 260–269, https://doi.org/10.4113/jom.2010.1061.

[3] M. Dodge, R. Kitchin, C. Perkins (Eds.), The Map Reader: Theories of Mapping Practice and Cartographic Representation, John Wiley & Sons, 2011.

[4] H.A. Taylor, Mapping the understanding of understanding maps, in: P. Shah, A. Miyake (Eds.), The Cambridge Handbook of Visuospatial Thinking, Cambridge University Press, New York, 2005, pp. 295–333.

[5] Department of Homeland Security, Budget-in-brief Fiscal year 2019, 10/13/19, www.dhs.gov/publication/fy-2019-budget-brief, 2017.

[6] M.K. Lindell, Communicating imminent risk, in: H. Rodríguez, J. Trainor, W. Donner (Eds.), Handbook of Disaster Research, Springer, New York, 2018, pp. 449–477.

[7] M.J. Crozier, J. McClure, J. Vercoe, M. Wilson, The effects of land zoning information on judgments about earthquake damage, Environ. Hazards 38 (2) (2006) 143–152.

[8] Y. Cao, B.J. Boruff, I.M. Mcneill, Is a picture worth a thousand words ? Evaluating the effectiveness of maps for delivering wild fire warning information, Int. J. Disaster Risk Reduct. 19 (2016) 179–196, https://doi.org/10.1016/j.ijdrr.2016.08.012.

[9] Y. Cao, B.J. Boruff, I.M. McNeill, The smoke is rising but where is the fire? Exploring effective online map design for wildfire warnings, Nat. Hazards 88 (3) (2017) 1473–1501, https://doi.org/10.1007/s11069-017-2929-9.

[10] L. Cheong, S. Bleisch, A. Kealy, K. Tolhurst, T. Wilkening, M. Duckham, Evaluating the impact of visualization of wildfire hazard upon decision-making under uncertainty, Int. J. Geogr. Inf. Sci. 30 (7) (2016) 1377–1404, https://doi.org/10.1080/13658816.2015.1131829.

[11] K. Haynes, J. Barclay, N. Pidgeon, Volcanic hazard communication using maps : an evaluation of their effectiveness, Bull. Volcanol. (2007), https://doi.org/10.1007/s00445-007-0124-7.

[12] M.A. Thompson, J.M. Lindsay, J.C. Gaillard, The influence of probabilistic volcanic hazard map properties on hazard communication, Appl. Volcanol. 4 (6) (2015) 1–24, https://doi.org/10.1186/s13617-015-0023-0.

[13] H.M. Bell, G.A. Tobin, Efficient and effective? The 100-year flood in the communication and perception of flood risk, Environ. Hazards 7 (4) (2007) 302–311, https://doi.org/10.1016/j.envhaz.2007.08.004.

[14] H. de Moel, J. van Alphen, J.C.J.H. Aerts, Flood maps in Europe -methods, availability and use, Nat. Hazards Earth Syst. Sci. 9 (2) (2009) 289–301, https://doi.org/10.5194/nhess-9-289-2009.

[15] M. Hagemeier-Klose, K. Wagner, Evaluation of flood hazard maps in print and web mapping services as information tools in flood risk communication, Nat. Hazards Earth Syst. Sci. 9 (2) (2009) 563–574, https://doi.org/10.5194/nhess-9-563-2009.

[16] E. Maidl, M. Buchecker, Raising risk preparedness by flood risk communication, Nat. Hazards Earth Syst. Sci. 15 (7) (2015) 1577–1595, https://doi.org/10.5194/nhess-15-1577-2015.

[17] M. van Kerkvoorde, W. Kellens, E. Verfaillie, K. Ooms, Evaluation of web maps for the communication of flood risks to the public in Europe, Int. J. Cartogr. 4 (1) (2018) 49–64.

[18] K.D. Ash, R.L. Schumann III, G.C. Bowser, Tornado warning trade-offs: evaluating choices for visually communicating risk, Weather Clim. Soc. 6 (2014) 104–118.

[19] M.A. Casteel, J.R. Downing, Assessing risk following a wireless emergency alert: are 90 characters enough? J. Homel. Secur. Emerg. Manag. (2015) https://doi.org/10.1515/jhsem-2015-0024.

[20] I. Jon, S.-K. Huang, M.K. Lindell, Perceptions and reactions to tornado warning polygons: would a gradient polygon be useful? Int. J. Disaster Risk Reduct. 30 (2018) 132–144, https://doi.org/10.1016/j.idrr.2018.01.035.

[21] I. Jon, S.-K. Huang, M.K. Lindell, Perceptions and expected immediate reactions to severe storm displays, Risk Anal. 39 (2019) 274–290, https://doi.org/10.1111/risa.12896.

[22] M.K. Lindell, S. Huang, H. Wei, C.D. Samuelson, Perceptions and expected immediate reactions to tornado warning polygons, Nat. Hazards 80 (2016) 683–707, https://doi.org/10.1007/s11069-015-1990-5.

[23] K. Sherman-Morris, M. Brown, Experiences of smithville, Mississippi residents with the 27 april 2011 tornado, Natl. Weather Digest 36 (2) (2012) 93–101.

[24] S. Arlikatti, M.K. Lindell, C.S. Prater, Y. Zhang, Risk area accuracy and hurricane evacuation expectations of coastal residents, Environ. Behav. 38 (2) (2006) 226–247, https://doi.org/10.1177/0013916505277603.

[25] J. Cox, D. House, M.K. Lindell, Visualizing uncertainty in predicted hurricane tracks, Int. J. Uncertain. Quantification 3 (2013) 143–156.

[26] L. Liu, A.P. Boone, I.T. Ruginski, L. Padilla, M. Hegarty, S.H. Creem-Regehr, W. B. Thompson, C. Yuksel, D.H. House, Uncertainty visualization by representative sampling from prediction ensembles, IEEE Trans. Visual. Comput. Graph. 23 (9) (2017) 2165–2178.

[27] L. Liu, L. Padilla, S.H. Creem-Regehr, D.H. House, Visualizing uncertain tropical cyclone predictions using representative samples from ensembles of forecast tracks, IEEE Trans. Visual. Comput. Graph. 25 (1) (2019) 882–891.

[28] R. Meyer, K. Broad, B. Orlove, N. Petrovic, Dynamic simulation as an approach to understanding hurricane risk response: insights from the stormview lab, Risk Anal. 33 (8) (2013) 1532–1552, https://doi.org/10.1111/j.1539-6924.2012.01935.

[29] L.M. Padilla, G. Hansen, I.T. Ruginski, H.S. Kramer, W.B. Thompson, S.H. Creem-Regehr, The influence of different graphical displays on nonexpert decision making under uncertainty, J. Exp. Psychol. Appl. 21 (1) (2015) 37–46.

[30] H. Wu, M.K. Lindell, C.S. Prater, C.D. Samuelson, Effects of track and threat information on judgments of hurricane strike probability, Risk Anal. 34 (6) (2014) 1025–1039, https://doi.org/10.1111/risa.12128.

[31] H.C. Wu, M.K. Lindell, C.S. Prater, Process tracing analysis of hurricane information displays, Risk Anal. 35 (12) (2015) 2202–2220, https://doi.org/10.1111/risa.12423.

[32] H.C. Wu, M.K. Lindell, C.S. Prater, Strike probability judgments and protective action recommendations in a dynamic hurricane tracking task, Nat. Hazards 79 (1) (2015) 355–380, https://doi.org/10.1007/s11069-015-1846-z.

[33] Y. Zhang, C.S. Prater, M.K. Lindell, Risk area accuracy and evacuation from hurricane bret, Nat. Hazards Rev. 5 (August) (2004) 115–120, 10.1061/(ASCE)1527-6988(2004)5:3(115).

[34] M. Hegarty, The cognitive science of visual-spatial displays: implications for design, Top. Cognit. Sci. 3 (3) (2011) 446–474.

[35] G.L. Allen, C.R. Miller Cowan, H. Power, Acquiring information from simple weather maps: influences of domain specific knowledge and general visual-spatial abilities, Learn. Indiv Differ 16 (4) (2006) 337–349.

[36] M.R. Endsley, D.G. Jones, Designing for Situation Awareness: an Approach to User-Centered Design, second ed., CRC Press, Boca Raton FL, 2012.

[37] R. Colom, M.J. Contreras, J. Botella, J. Santacreu, Vehicles of spatial ability, Pers. Indiv. Differ. 32 (5) (2002) 903–912.

[38] B. Aksoy, Investigation of mapping skills of pre-service teachers as regards to various parameters, Educ. Res. Rev. 8 (4) (2013) 134–143, https://doi.org/10.5897/ERR12.221.

[39] G. Albert, V. Ilyés, D. Kis, C. Szigeti, D. Várkonyi, Testing the map reading skills of university students, in: 6th International Conference on Cartography and GIS, 2016, pp. 188–199.

[40] A.J. Milson, M. Alibrandi, Critical map literacy and geographic information systems: the spatial dimension of civic decision making, in: P.J. VanFossen, M. J. Berson (Eds.), The Electronic Republic? the Impact of Technology on Education for Citizenship, Purdue University Press, West Lafayette, 2008, pp. 110–128.

[41] S.P. Muir, Understanding and improving students' map reading skills, Elem. Sch. J. 86 (2) (1985) 206–216.

[42] S.M. Kosslyn, Understanding charts and graphs, Appl. Cognit. Psychol. 3 (1989) 185–226.

[43] P. Shah, E.G. Freedman, I. Vekiri, The comprehension of quantitative information in graphical displays, in: P. Shah, A. Miyake (Eds.), The Cambridge Handbook of Visuospatial Thinking, Cambridge University Press, New York, 2005, pp. 426–476.

[44] M. Kunz, L. Hurni, How to enhance cartographic visualisations of natural hazards assessment results, Cartogr. J. 48 (1) (2011) 60–71, https://doi.org/10.1179/1743277411Y.0000000001.

[45] C.A. Brewer, Designing Better Maps, first ed., ESRI Press, Redlands, CA, 2005.

[46] M. Harrower, C.A. Brewer, ColorBrewer.org: an online tool for selecting color schemes for maps, Cartogr. J. 40 (1) (2003) 27–37.

[47] M.K. Lindell, R.W. Perry, The protective action decision model: theoretical modifications and additional evidence, Risk Anal. 32 (4) (2012) 616–632, https://doi.org/10.1111/j.1539-6924.2011.01647.x.

[48] C. Keller, Using a familiar risk comparison within a risk ladder to improve risk understanding by low numerates: a study of visual attention, Risk Anal. 31 (7) (2011) 1043–1054.

[49] C. Kreuzmair, M. Siegrist, C. Keller, High numerates count icons and low numerates process large areas in pictographs: results of an eye-tracking study, Risk Anal. 36 (8) (2016) 1599–1614.

[50] J.B. Carroll, Human Cognitive Abilities: A Survey of Factor Analytic Studies, Cambridge University Press, Cambridge, 1993.

[51] M.G. McGee, Human spatial abilities: psychometric studies and environmental, genetic, hormonal, and neurological influences, Psychol. Bull. 86 (5) (1979) 889–918.

[52] R.B. Ekstrom, J.W. French, H.H. Harman, D. Dermen, Manual for Kit of Factor-Referenced Cognitive Tests, Educational Testing Service, Princeton, NJ, 1976.

[53] M. Hegarty, R.D. Crookes, D. Dara-Abrams, T.F. Shipley, Do all science disciplines rely on spatial abilities? Preliminary evidence from self-report questionnaires, in: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics),, vol. 6222, LNAI, 2010, pp. 85–94, https://doi.org/10.1007/978-3-642-14749-4_10.

[54] K. Ooms, P. De Maeyer, V. Fack, Listen to the map user: cognition, memory, and expertise, Cartogr. J. 52 (1) (2015) 3–19, https://doi.org/10.1179/1743277413Y.0000000068.

[55] Y. Postigo, J.I. Pozo, On the Road to Graphicacy: the learning of graphical representation systems, Educ. Psychol. 24 (5) (2004) 623–644, https://doi.org/10.1080/0144341042000262944.

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

[57] A. Bosco, A.M. Longoni, T. Vecchi, Gender effects in spatial orientation: cognitive profiles and mental strategies, Appl. Cognit. Psychol. 18 (5) (2004) 519–532.

[58] P.W. Thorndyke, C. Stasz, Individual differences in procedures for knowledge acquisition from maps, Cognit. Psychol. 12 (1980) 137–175.

[59] S.G. Vandenberg, A.R. Kuse, Mental rotations, A group test of three-dimensional spatial visualization, Percept. Mot. Skills 47 (1978) 599–604.

[60] M. Campos, C. Rica, R.W. Uk, Key Vulnerabilities Coordinating Lead Authors : Contributing Authors : Review Editors, 2014, pp. 1039–1099.

[61] K. Kastens, Commentary: object and spatial visualization in geosciences, J. Geosci. Educ. 58 (2) (2010) 52–57, https://doi.org/10.5408/1.3534847.

[62] R.E. Mayer, L.J. Massa, Three facets of visual and verbal learners: cognitive ability, cognitive style, and learning preference, J. Educ. Psychol. 95 (4) (2003) 833–846, https://doi.org/10.1037/0022-0663.95.4.833.

[63] A. Nazareth, N.S. Newcombe, T.F. Shipley, M. Velazquez, S.M. Weisberg, Beyond small-scale spatial skills: navigation skills and geoscience education, Cognit. Res.: Princ. Implications 4 (1) (2019), https://doi.org/10.1186/s41235-019-0167-2.

[64] L. Tartre, Spatial orientation skill and mathematical problem solving, J. Res. Math. Educ. 21 (3) (1990) 216–229.

[65] S.M. Weisberg, V.R. Schinazi, N.S. Newcombe, T.F. Shipley, R.A. Epstein, Variations in cognitive maps: understanding individual differences in navigation, J. Exp. Psychol. Learn. Mem. Cognit. 40 (3) (2014) 669–682, https://doi.org/10.1037/a0035261.

[66] C.B. McCormick, Metacognition and learning, in: W.M. Reynolds, G.E. Miller (Eds.), Handbook of Psychology: Educational Psychology, vol. 7, John Wiley & Sons Inc., 2003, pp. 79–102.

[67] D. Dunning, The Dunning–Kruger effect: on being ignorant of one's own ignorance, in: J.M. Olson, M.P. Zanna (Eds.), Advances in Experimental Social Psychology, vol. 44, Academic Press, 2011, pp. 247–296.

[68] M. Hegarty, A.E. Richardsona, D.R. Montello, K. Lovelace, I. Subbiah, Development of a self-report measure of environmental spatial ability, Intelligence 30 (5) (2002) 425–447, https://doi.org/10.1016/S0160-2896(02)00116-2.

[69] R.L. Ott, M. Longnecker, An Introduction to Statistical Methods and Data Analysis, sixth ed., Duxbury, Pacific Grove CA, 2010.

[70] J.C. Nunnally, I.H. Bernstein, Psychometric Theory, third ed., McGraw-Hill, New York, 1994.

[71] R. Fall, N.M. Webb, N. Chudowsky, Group discussion and large-scale language arts assessment: effects on students' comprehension, Am. Educ. Res. 38 (4) (2000) 911–941.

[72] L.M. Padilla, I.T. Ruginski, S.H. Creem-Regehr, Effects of ensemble and summary displays on interpretations of geospatial uncertainty data, Cognit. Res.: Princ. Implications 2 (1) (2017) 40.