

# Evaluating the Effect of Display Realism on Natural Resource Decision Making

Steven S. Chong<sup>a</sup>

<sup>a</sup> School of Information, University of Arizona, Tucson, Arizona, United States; [stevenchong@email.arizona.edu](mailto:stevenchong@email.arizona.edu)

**Abstract:** Geographic information systems (GIS) facilitate location-based decision making. Despite the improved availability of GIS software to non-professionals, training in cartographic design has not followed suit. Prior research indicates that when presented with map choices, users are influenced by naïve realism, a preference for realistic displays containing irrelevant, extraneous details, leading to decreased task efficiency. This study investigated the role of naïve realism in decision making for natural resource management, a field that often employs geospatial tools. Data was collected through a GIS user ability test, a questionnaire and direct observation. Forty volunteer expert and non-expert resource managers evaluated the suitability of different sites for a land management scenario. Each participant was tested on two map display treatments containing different levels of realism – a simpler 2D display and a more complex 3D display – to compare task performance. Performance was measured by task accuracy and task completion time. User perceptions and preferences about the displays were also recorded. Display realism had an impact on performance and there were indications naïve realism was present. Users completed tasks significantly faster on the 2D display and many individuals misjudged which display they were most accurate or fastest with. The results are informative for designing information systems containing interactive maps, particularly for resource management applications. The results also suggest that the order displays were presented had a significant effect and may have implications for teaching users map-based tasks.

**Keywords:** GIS, User study, Cartography, Cognitive study, User interactions

## 1. Introduction

Geographic information systems (GIS) and dynamic maps facilitate decision making. Despite it being easier for novice users without formal cartographic training to create maps, they often do not choose the most effective visual displays for completing tasks, leading to increased cognitive load and decreased performance.

A user's experience level affects how cartographic information is accessed and used. For example, Chang et al. (1985) found a difference in how experts and novices interpreted topographic maps. Expert map readers were better able to recognize landform patterns and determine the high and low points on test maps because they searched for patterns in contour lines to match with familiar landform patterns stored in their memory. This allowed experts to quickly associate the two-dimensional (2D) map to a three-dimensional (3D) representation. Other studies relevant to how users interact with GIS have indicated that users with a higher level of GIS experience tend to access and display data in a more systematic manner than less experienced users (McGuinness et al. 1993; McGuinness 1994). This research indicates that experience level influences what strategies are employed to access and use cartographic information. Experts made the most use of available information by examining fewer variables at the same time, while also reviewing more combinations of variables than novices.

Visual presentation of information is not trivial and has an impact on the decision making process. McKendry (2000) found that by varying graphic organization and

breaking normally accepted rules for presenting relationships on maps the difficulty of completing a site selection task increased as the quality of the display decreased. Prior studies found users prefer and have more confidence in realistic displays than when using simpler, more abstract displays, despite the extra time and effort required for completing tasks (Hegarty 2011; Smallman and Cook 2011). It was found that more detailed displays hinder task completion by showing variables that are irrelevant to the completion of assigned tasks. The disconnect between the user's preference for realism and the user's lack of awareness about its negative effect on task performance is termed naïve realism by Smallman and St. John (2005).

It is unclear to what degree naïve realism affects task completion and decision performance with interactive maps. The majority of studies involved static displays with relatively simple tasks, including map reading and inference. Hegarty et al. (2009) study found meteorologists use configurable displays but only questioned participants about their preferences and did not examine performance. Other studies have had users place routes on terrain maps (Smallman and Cook 2011; Smallman et al. 2007; Smallman and St. John 2005) or find relationships in static maps containing extraneous variables (Hegarty 2013; Hegarty et al. 2012; Hegarty et al. 2009).

This study investigated the role of naïve realism in decision making for natural resource management. Resource management makes extensive use of GIS for supporting decision making, such as determining the

range of invasive species or developing habitat suitability models. Data were collected from users through a GIS user ability test, direct observation and a questionnaire to ascertain if display realism affects performance on site selection tasks. Re-al-world field data were used to develop the scenario and GIS for user testing. Site selection is a common decision making task on a GIS used in economic, urban, and ecological analyses and comparisons of user performance on a set of rule-based site selection tasks were used to determine if there was any relationship between domain experience, display realism, and task performance.

## 2. Data

### 2.1 Scenario Datasets

In this study, participants assumed the roles as land managers for a hypothetical scenario. Users interpreted GIS datasets containing vegetation, roads, slope, and precipitation data and site locations. The datasets were imported into the GIS software as separate layers and were the criteria that users had to base their decision making on. Re-source management professionals at the University of Arizona and Luke Air Force Base provided the datasets and validated the GIS visualizations, scenario and site selection tasks for the user test. A pilot study and pretest con-firmed that users without any formal training in natural resource management or GIS were able to interpret the datasets correctly.

### 2.2 Display Treatments

The GIS displays for the user tests were created with ESRI's ArcGIS 10.3 software and its 3D Analyst extension. Display realism was represented at two levels, 2D and 3D. Both display treatments contained the same GIS layers. The 2D display had a simple base map that depicted terrain information in 2D. The 3D display contained a 30-meter resolution digital elevation model (DEM) and had high resolution aerial imagery draped over it to further enhance realism. Figures 1 and 2 illustrate example screen captures of the precipitation layer for the 2D and 3D dis-plays treatments, respectively. It should be noted that elevation data were not required to evaluate the site selection criteria in this study.

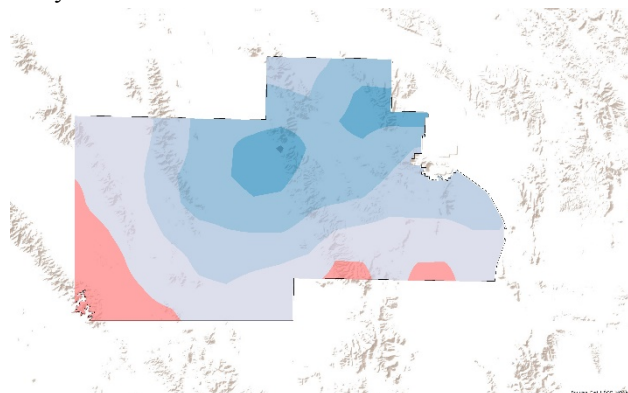


Fig. 1. Screen capture of the precipitation layer in the 2D display.

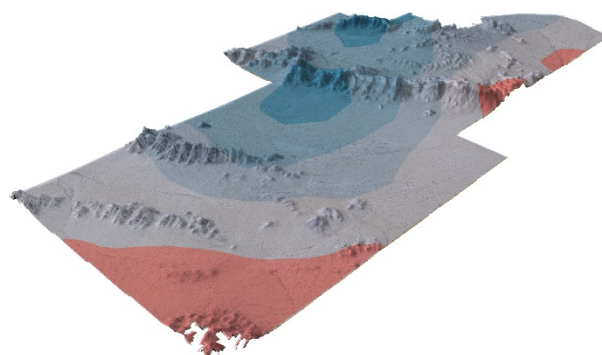


Fig. 2. Screen capture of the precipitation layer in the 3D display.

### 2.3 Methodology

The study utilized a user ability test, direct observation, and a post-test questionnaire to assess user performance, decision confidence, preferences, intuitions about performance, and demographic information. For the user tests, participants used a spreadsheet to rank alternate sites by following provided guidelines and analyzing the GIS dis-plays, similar to the approach taken by Mennecke et al. (2000) in the business domain. Figure 3 shows a copy of the spreadsheet that users filled out during the user test. Performance on site selection tasks is commonly evaluated by task completion time and accuracy (Erskine et al. 2013). Research on naïve realism often uses the same metrics to assess map reading and interpretation tasks (e.g. Hegarty et al. 2012; Hegarty et al. 2009). Both were therefore selected to evaluate task performance in this study.

Decision Scoring Sheet						
1) Enter your study ID number in the green area						
2) Use the GIS and guidelines to fill in all the cells highlighted in yellow. The total scores are calculated automatically. After all scores are entered, use the total scores to rank the sites from 1 to 5 and enter the rankings in the blue area.						
Criteria	Guidelines	Sites				
		A	B	C	D	E
Distance from Roads	If site greater than 5 km from a road, 5 points					
	If site 1 km to 5 km from a road, 3 points					
	If site less than 1 km from a road, 0 points					
Slope Percentage	If site located on 0% to 5% slope, 5 points					
	If site located on 5% to 20% slope, 3 points					
	If site located on 20% or greater slope, 1 point					
Precipitation - Difference from Normal	If site located in area with precipitation 2 or more inches above normal, 5 points					
	If site located in area with precipitation 0 to 2 inches above normal, 3 points					
	If site located in area with below normal precipitation, 1 point					
Vegetation Association	If site located in a creosote-white bursage association, 5 points					
	If site located in any other creosote-bursage association, 4 points					
	If site located in any other association, 1 point					
Site Score Totals		0	0	0	0	0
Site Ranking (higher scores are better; 1 = best, ties are okay)						
3)						
4) On a scale of 1 to 5, how confident are you about your decisions? (1 = low confidence, 5 = high confidence)						

Fig. 3. Screen capture of a spreadsheet containing the guidelines users followed for the site selection tasks.

### 2.4 Participants

The participants were 40 students at the University of Arizona (20 experts, 20 novices; 19 females, 21 males) and they were separated into expert and novice user groups. The expert participants were graduate students in natural resource management and related departments. These students each had years of experience in resource management and using GIS. The novice group was composed of undergraduate students with no experience in re-source management or using GIS. The participants were volunteers and prescreened for colorblindness.

Monetary compensation was provided based on participation and performance.

## 2.5 Procedure

The user tests took place in an office at the University of Arizona over the course of a month during an academic semester. Testing was conducted on an individual basis with the researcher serving as a proctor. Participants received an orientation to familiarize themselves with the scenario, using GIS software, the scoring sheet, and the goals of the exercise. Time was provided for users to explore using the GIS software firsthand on test data. Users were instructed to not use any GIS analysis tools and functions aside from those introduced during the orientation. The user ability test was counterbalanced, so that half of the participants received the 2D display first. After the user test concluded, each participant completed the questionnaire.

## 3. Results and Discussion

SPSS software (version 23) was used to perform statistical analyses and an alpha level of .05 was used to determine the significance of the statistical tests. Two users had their scores omitted due to encountering software errors during their tasks.

### 3.1 User Accuracy

User accuracy was calculated with the approach taken by Mennecke et al. (2000) by measuring Kendall's Tau coefficient to compare the participant's site rankings to an answer sheet. Accuracy rates were evaluated with paired samples sign tests for within-group analyses and Mann Whitney U tests for comparisons between groups. No significant differences in accuracy rates were found between tasks performed on the 2D and 3D displays. In general, participants scored consistently high accuracy rates with a mean of 92.0 percent on tasks with the 2D display ( $n = 38$ ,  $\bar{x} = .920$ ,  $s = .155$ ) and 96.0 percent for tasks on the 3D display ( $n = 38$ ,  $\bar{x} = .960$ ,  $s = .099$ ). Both experts and novices were highly accurate on both display types. Experts had a mean score of 93.0 percent on the 2D display tasks ( $n = 19$ ,  $\bar{x} = .930$ ,  $s = .179$ ) and 95.3 percent on the 3D display tasks ( $n = 19$ ,  $\bar{x} = .953$ ,  $s = .120$ ). Similarly, novices had a mean score of 91.1 percent on the 2D display tasks ( $n = 19$ ,  $\bar{x} = .911$ ,  $s = .131$ ) and 96.7 percent on the 3D display tasks ( $n = 19$ ,  $\bar{x} = .967$ ,  $s = .075$ ). Accuracy performance in females and males were also explored. In general, male users scored higher mean 2D accuracy scores ( $n = 20$ ,  $\bar{x} = .943$ ,  $s = .160$ ) than female users ( $n = 18$ ,  $\bar{x} = .895$ ,  $s = .150$ ). Males also scored higher on 3D display accuracy ( $n = 20$ ,  $\bar{x} = .977$ ,  $s = .062$ ) than female users ( $n = 18$ ,  $\bar{x} = .942$ ,  $s = .128$ ).

It was predicted that increased dimensionality, in addition to the extraneous aerial imagery, would have increased the difficulty of tasks on the 3D display. It was also expected expert users familiar with GIS would have used their prior experience to compensate for the extraneous details, such as turning off extra layers, while novices would not, thereby reducing user accuracy for novice users that did not have experience using GIS. Male users

were also expected to score higher than females, as they tend to perform better on spatial perception and mental rotation tests (Linn & Petersen, 1985). Although males, in general, did have higher accuracy scores than female users, there was not a significant difference for either display type.

There are several possible explanations for the negligible difference in the accuracy scores. The fact that users, in general, scored a mean accuracy rate over 90 percent indicates the majority of users acquired relevant knowledge and performed well on both site selection tasks on either display type. Many achieved scores of 100 percent, which resulted in strong negative skewing of the accuracy results. The large number of high scoring participants likely accounted for the insignificant differences in accuracy scores within the subgroups. Users were also not limited by the amount of time spent on the tasks. Although users were instructed to complete their tasks as fast as possible, they received a financial bonus for submitting correct answers. This may have motivated participants to check their work and likely increased overall accuracy rates.

### 3.2 User Completion Times

The user ability tests were timed and on average, participants completed the 2D display task in 700.53 seconds ( $n = 38$ ,  $\bar{x} = 700.530$ ,  $s = 257.852$ ) and 3D display task in 833.63 seconds ( $n = 38$ ,  $\bar{x} = 833.630$ ,  $s = 402.682$ ), indicating users required more time to complete tasks on the realistic display. Adding details to maps can be helpful as long as the details are relevant to intended tasks. If not, the additional details may hinder task completion because the useful information is obscured and it takes longer for users to extract relevant information. With regards to this user study, the 3D maps contained elevation information and aerial imagery that were unnecessary for completing the site selection tasks. Users working with the 3D display may have had certain perspectives obscured by the terrain, requiring users to rotate the map or change the zoom level. This interactivity would cost users time, requiring them to change perspectives. In contrast, with the 2D map users saw the entire study area from a high-level view and did not have to rotate the display to evaluate all of the sites.

Comparing the different expertise groups showed that experts had a mean completion time of 647.840 seconds ( $n = 19$ ,  $\bar{x} = 647.840$ ,  $s = 284.117$ ), while novices completed tasks in a mean of 753.210 seconds in the 2D display ( $n = 19$ ,  $\bar{x} = 753.210$ ,  $s = 223.801$ ). Regarding tasks on the 3D display, experts completed tasks in a mean of 766.210 seconds ( $n = 19$ ,  $\bar{x} = 766.210$ ,  $s = 508.237$ ) and novices finished on average in 901.050 seconds ( $n = 19$ ,  $\bar{x} = 901.050$ ,  $s = 255.762$ ). Although experts completed tasks faster than novices on both displays, a comparison of expert and novice completion times indicated there was only a marginally significant difference between expert and novice completion times on the 3D display ( $t(36) = -2.012$ ,  $p = .052$ ). However, within-groups comparisons found that novices were much faster on the 2D display while the expert users did not

perform differently ( $t(36) = -1.663, p = .105$ ). These results suggest that novice users were more affected by dimensionality – performing tasks on the 3D display required novices to spend more time on the tasks.

It was expected expert users would perform well with tasks on the 2D display because the user test employed common tasks on a GIS that should not have been difficult for those with GIS experience – having a minimum of one year's worth of experience was a requirement for participation. Having prior GIS experience appeared to benefit many expert users as they employed the GIS tools more efficiently than the novice group. Experts tended to use tools more than novice users and this was likely due to their familiarity with using them. By appropriately using the GIS tools, expert users may have negated some of the visual challenges the 3D display posed and employed the tools as a form of compensation, as described by McKendry (2000).

Another explanation for experts having similar results for completion times in both displays is that many experts minimized the number of rotations made in the 3D display and essentially treated it as a 2D display. These users often rotated the 3D map so that it would show a top-down overview of the study area that was akin to the view shown by the 2D display. Minimizing the number of map rotations while simultaneously evaluating each site would have led to faster task completion than having to rotate the display and changing perspectives to view each site. The novice users tended to rotate the map often during the 3D display tasks and this likely accounted for some of the additional time required to complete the tasks over the ones on the 2D display.

An exploration of female and male completion times found that females were significantly faster when using the 2D display with a mean completion time of 727.940 seconds ( $n = 18, \bar{x} = 727.940, s = 285.342$ ) than the 3D display with a mean completion time of 838.390 seconds ( $n = 18, \bar{x} = 838.390, s = 285.131$ ). In contrast, male users did not show any differences when performing tasks on either display and completed tasks on the 2D display in an average of 675.85 seconds ( $n = 20, \bar{x} = 675.850, s = 235.103$ ) and on the 3D display with a mean time of 829.350 seconds ( $n = 20, \bar{x} = 829.350, s = 492.939$ ). There were no significant differences between female and male completion times for either display. These results suggest that sex did not have any major effect on completion times for this study. However, females were faster with 2D tasks than 3D tasks, suggesting they were more negatively affected by display realism. This may have been due to the 3D display requiring users to rotate map perspectives to interpret the values for variables in each site – according to prior research females tend to have more difficulty with mental rotation and spatial perception tasks compared to males (Linn & Petersen, 1985).

The data were analyzed with a multivariate analysis of covariance (MANCOVA) to examine if any relationships existed between the independent variables and completion times for both display types. There were

some indications that display order affected completion time, therefore, the first display presented to users was included in this analysis. The independent variables were Expertise, Sex, and First Display Presented. Covariates were included in the analysis, such as Age, Natural Resource Management Experience, GIS Experience, Experience with 3D Software, GIS Tool Use (2D Display), GIS Tool Use (3D Display). The MANCOVA results revealed a significant multi-variate main effect for First Display Presented (Wilks'  $\lambda = .382, F(2, 23) = 18.622, p < .001, \text{partial } \eta^2 = .618$ ). No significant interactions between variables were detected. After accounting for a Bonferroni correction, none of the univariate ANOVA results indicated a significant effect.

### 3.3 Display Order

User accuracy was compared for users receiving the 2D display treatment first versus those receiving the 3D display first. No significant difference in accuracy scores was found. In contrast, the completion time results indicate that display order had a significant effect ( $t(17) = -6.509, p < .001$ ). Further analysis of total user completion times revealed that users receiving the 2D display treatment first completed all tasks in a mean of 1568.95 seconds, and those receiving the 3D display first completed all tasks in a total of 1495.5 seconds, on average. In other words, users receiving the 3D treatment first were 73.45 seconds faster, on average. Users receiving the 3D display first completed their subsequent 2D display tasks significantly faster – approximately over 330 seconds (5.5 minutes) faster, on average.

Taken altogether, the results indicate there was a greater learning effect from users receiving the 3D display treatment first. It appears that presenting users with the 3D display first allowed them to learn the site selection tasks better and perform similar tasks much faster on the simpler 2D display. These results are supported by Vessey's (1991) work with graph visualizations that found memorization improved when the same information was presented in a 3D format over a 2D format. Those that received the 3D display treatment first spent more time navigating the map and likely had to concentrate on their tasks to score high accuracy. Having the ability to better memorize information related to the tasks would have resulted in much faster completion times during the second treatment.

### 3.4 Naive realism, User Preferences, and User Confidence

The questionnaire data was used to measure user perceptions about accuracy rates and completion times and determine if any users were susceptible to naive realism. Both the expert and novice user groups had individuals exhibiting naive realism for completion times and accuracy.

The questionnaire data was analyzed to determine how users believed they performed in terms of speed. Although the majority of users (29 out of 38) believed they were faster with the 2D display, 9 users thought they

were faster on the 3D display or performed equally fast on both displays. Of these 9 users, 7 were faster with the 2D display, indicating the presence of naïve realism (6 users perceived no difference in performance times but actually performed faster on the 2D display). These naïve users completed tasks 291 seconds faster, on average, with the 2D display. The results indicate that novice users had more difficulty making accurate time judgments, as the naïve group had 2 experts and 5 novices. In terms of sex, there were 3 female and 4 male naïve users. Each of the naïve users received the 3D display treatment first.

The questionnaire data was also examined to determine users' perceptions about task accuracy. Although only 34.2% of all users thought they were most accurate with the 3D display (16 out of 38 users), they outnumbered those that preferred the 2D display (11 out of 38 users) and 11 users were indifferent. Of the 16 users preferring the 3D display, 13 were actually more accurate with the 2D display and exhibited naïve realism. The majority of these users were in the expert group (11 users), and users were almost evenly divided based on sex (6 females and 7 males). Nine of the naïve users received the 2D treatment first.

The majority of users had difficulty judging their own accuracy. Only 12 users correctly reported the display they performed most accurately with. In contrast, 20 users correctly reported the display they were fastest with. The results indicate users were better at judging their completion time performance than their accuracy performance, with experts having more errors perceiving the latter. This is likely due to experts believing that having any additional information aids with decision making even though the extra information may not be relevant to the decision to be made. The phenomenon observed echoes prior research in naïve realism (Hegarty et al. 2012; Hegarty et al. 2009). Many of the experts reported in their questionnaire responses that the 3D elevation data was helpful, when it was not useful for the tasks assigned in this study. In this study, novices do not seem to be as eager for additional information to complete the tasks, either because they lack subject knowledge, causing them to realize that additional information was unnecessary or they were not as engaged or invested in the goals of the task, as the experts were. Many of the expert users were careful with their tasks and checked their answers. The expert users were recruited from disciplines involving natural resources and were likely more interested in the scenario topic than the novices, possibly causing the experts to overthink the tasks. No differences in user accuracy were detected between the expertise groups and this may be due to most of the users scoring highly on the user tests. Users who were naïve about their time performance received the 3D treatment first and thought they did better with that display even though they were faster with the 2D display. It may be that because users were so engrossed with becoming familiar with the task during their first treatment they did not realize how much time had passed. Similarly, the majority of users naïve about their accuracy

performance tended to have the 2D display treatment first. These users thought they improved their performance in the second treatment when they actually performed just as well on both displays.

User confidence data was gathered at the end of each site selection task. No significant difference in confidence levels was found within the expert or novice groups, or between groups. These results run counter to those in Fabrikant and Boughman (2006) and Zanola et al. (2009), where significant differences in confidence levels between experts and novices were found. However, those studies involved static maps and interacting with the 3D maps in this study may have led users to question their confidence levels because it was the first time many participants used the software. It also may be the case that user familiarity with 2D maps caused them to feel more confident using that display. Sixteen users reported liking the 3D display best and the majority of these users listed reasons including it allowing better visualization of the terrain and it being more realistic. It should be noted that eight of these users were naïve about accuracy because they scored equally well with both the 2D and 3D tasks but believed they performed better with the 3D display.

#### 4. Conclusion

This study explored how display realism affects decision making performance with interactive maps and user perceptions about their own performance. Cartographic visualizations that allow users to view maps with a third dimension are becoming more accessible to users. Prior research with static maps laid down a foundation for a natural progression to study task performance with interactive maps, allowing comparisons of maps with varying levels of realism. Users appeared to have greater difficulty evaluating the accuracy of their decisions and a larger percentage of users were influenced by naïve realism for accuracy than for completion time. The results indicate that for instances where map-based task efficiency or accuracy are of the utmost importance, it may be prudent to limit users' choice in map displays as many users had trouble judging their own performance and preferred the more realistic 3D display.

This study also found that expert and novice users took different approaches during the user tests. It was observed that experts generally followed an approach for the site selection tasks that minimized the number of map rotations and employed GIS tools. These factors likely reduced the negative effects of display realism when compared to novice user performance, particularly with regards to completion time. The results indicate that novices required significantly more time to complete tasks on the realistic display when compared to the expert group. These results may also aid the development of customized interfaces for users with varying levels of expertise; novice users with little GIS and natural resource experience, for example, may benefit by having their map choices limited to simpler displays to speed up task performance.



The order that the display treatments were presented to users had a significant impact on task performance. It appears that users learned to perform tasks faster when presented with the 3D display first, followed by the 2D display. These results may have implications for teaching users how to perform map-based tasks more efficiently. The results suggest it may be more effective to train users on 3D displays and perform subsequent tasks on 2D displays. With this method, users would benefit from the increased memorization that 3D visualizations facilitate and become less distracted with the simpler 2D display. This method may particularly be useful in professions where task completion time is critical, such as emergency response. However, it should be emphasized that this study involved a specific set of tasks for natural resource management and additional research is necessary to confirm the applicability of these results to other map-based tasks and disciplines. However, the results do hold promise for stimulating future research that may improve map-based decision making performance.

One potential application of this research is assessing the impact of cartographic display realism on disaster management. Chen et al. (2016) have developed a methodology for creating evacuation maps in response to hurricanes and addressed the importance of cartographic elements, such as color and symbol choice, to communicate hazard risk to the public; choice of base map or level of map realism were not discussed in the article.

As technology advances, the implications of this study's results may also extend to new visualization techniques. For example, CAVE (cave automatic virtual environment) technology is increasingly being used for geospatial applications in immersive environments (Batty et al. 2017). The consequences of employing virtual reality technology are largely unknown for map-based task performance. It would be worthwhile to investigate if the immersive nature of these new technologies facilitate or have any detrimental result on task performance. Cartographers have a role in guiding users with the application of these new technologies – they could best train users how to make the most of the increased access to mapmaking technologies and improve map designs for better task performance.

## 5. Acknowledgements

The author would like to thank the School of Information at the University of Arizona for its funding support during the development of this work.

## 6. References

Batty, M., Lin, H., & Chen, M. (2017). 7. Virtual realities, analogies and technologies in geography. *Handbook on Geographies of Technology*, 96.

Chang, K. T., Antes, J., & Lenzen, T. (1985). The effect of experience on reading topographic relief information: Analyses of performance and eye movements. *The Cartographic Journal*, 22(2), 88-94.

Chen, Y. H., Zick, S. E., & Benjamin, A. R. (2016). A comprehensive cartographic approach to evacuation map creation for Hurricane Ike in Galveston County, Texas. *Cartography and Geographic Information Science*, 43(1), 68-85.

Erskine, M. A., Gregg, D. G., Karimi, J., & Scott, J. E. (2013). Business decision-making using geospatial data: A research framework and literature review. *Axioms*, 3(1), 10-30.

Fabrikant, S.I. & Boughman, A., 2006. Communicating Data Quality through Realism. In *Proceedings of GIScience 2006 (Extended Abstracts)*. Münster, Germany, pp. 59–60.

Hegarty, M. (2011). The cognitive science of visual-spatial displays: Implications for design. *Topics in cognitive science*, 3(3), 446-474

Hegarty, M. (2013). Cognition, metacognition, and the design of maps. *Current Directions in Psychological Science*, 22(1), 3-9.

Hegarty, M., Smallman, H. S., & Stull, A. T. (2012). Choosing and using geospatial displays: Effects of design on performance and metacognition. *Journal of Experimental Psychology: Applied*, 18(1), 1-17.

Hegarty, M., Smallman, H. S., Stull, A. T., & Canham, M. S. (2009). Naïve cartography: How intuitions about display configuration can hurt performance. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 44(3), 171-186.

Linn, M. C., & Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child development*, 1479-1498.

McGuinness, C., Van Wersch, A., & Stringer, P. (1993). User differences in a GIS environment: A protocol study. *Proceedings, 16th Conference of the International Cartographic Association*, May 3-9, Cologne, Germany, pp. 478-485.

McGuinness, C. (1994). Expert/novice use of visualization tools. In A. MacEachren & D. Taylor (Eds.), *Visualization in modern cartography* (Vol. 2, pp. 185-199). Oxford, England: Pergammon.

McKendry, J. (2000). The influence of map design on resource management decision making. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 37(2), 13-27.

Mennecke, B. E., Crossland, M. D., & Killingsworth, B. L. (2000). Is a map more than a picture? The role of SDSS technology, subject characteristics, and problem complexity on map reading and problem solving. *MIS Quarterly*, 601-629.

Smallman, H. S., & Cook, M. B. (2011). Naïve Realism: Folk fallacies in the design and use of visual displays. *Topics in cognitive science*, 3(3), 579-608.

Smallman, H. S., Cook, M. B., Manes, D. I., & Cowen, M. B. (2007, October). Naïve realism in terrain appreciation. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 51, No. 19, pp. 1317-1321). Sage Publications.

Vessey, I. (1991). Cognitive fit: A theory-based analysis of the graphs versus tables literature. *Decision Sciences*, 22(2), 219-240.

Zanola, S., Fabrikant, S. I., & Çöltekin, A. (2009, November). The effect of realism on the confidence in spatial data quality in stereoscopic 3D displays. In *Proceedings of the 24th international cartography conference (ICC 2009)*, Santiago, Chile.