Examining Training Comprehension and External Cognition in Evaluations of Uncertainty Visualizations to Support Decision Making

Jihye Song, Olivia B. Newton, Stephen M. Fiore, Corey Pittman, & Joseph J. LaViola, Jr.
University of Central Florida

Recent advances in uncertainty visualization research have focused not only on design features to support decision making, but also on challenges of evaluating the effectiveness of uncertainty visualizations, such as the degree to which individuals' baseline task comprehension may alter their performance on experimental tasks regardless of a visualization's effectiveness. Building on recent work, we investigated the effect of training comprehension on performance across varying representations of uncertainty and varying degrees of visualization interactivity using a simulated course of action selection task. Additionally, we explored how extended cognition theory can be applied to visualization evaluations by incorporating interface features that afford externalization of knowledge within the task environment. Our findings suggest that regardless of how uncertainty is represented, training comprehension leads to superior transfer, reduced workload, more accurate metacognitive judgments, and higher cognitive efficiency. Our findings also suggest that external cognition during decision making leads to improved accuracy and cognitive efficiency. The present study contributes to research on the design and evaluation of uncertainty visualizations. In addition, this study extends previous work by demonstrating how extended cognition theory can inform the design of human-machine interfaces to support decision making.

Visualizations to support decision making under uncertainty have been a topic of significant interest (Kinkeldey, MacEachren, Riveiro, & Schiewe, 2017), with implications for operations in complex environments. Recent advances in uncertainty visualization research have focused not only on design features to support decision making, but also the degree to which individuals' baseline task comprehension may alter performance regardless of a visualization's effectiveness. In particular, recent work has identified challenges in uncertainty visualization research with non-expert participants. Specifically, due to the complex nature of decision-making tasks, combined with participants' lack of experience with both the task itself and uncertainty visualization interfaces, participants may not adequately understand how to complete the task. This can lead to impaired performance, which could then obfuscate results by suggesting the visualizations are ineffective.

The present study builds upon recent work investigating factors that alter performance on decision-making tasks in empirical assessments of uncertainty visualizations. In particular, we focus on the importance of understanding the degree to which training alters performance, as well as ways to support decision making through interactive interfaces.

 DOI 10.1177/1071181319631520

Uncertainty Visualization Challenges

Despite the amount of interest in understanding how to design effective visualizations of uncertainty, a potential limitation of empirical evaluations in this area concerns the inherent difficulty non-expert participants face when they engage in tasks requiring domain-specific knowledge. For instance, prior work examining this limitation has shown that performance on a decision-making task utilizing uncertainty visualizations may vary not just due to the effectiveness of the visualization, but also as a function of how well participants comprehend the task (Fiore et al., 2019; Song et al., 2018).

In the extant literature, a significant amount of research has identified differences in how experts and non-experts perform on complex operational tasks. These differences present a challenge for uncertainty visualization research, in that visualizations may be designed with expert users in mind, but evaluated with participants who lack specialized domain knowledge and relevant experience. Some studies have approached this issue through efforts like modifying experimental tasks to facilitate comprehension by non-experts (Kirschenbaum, Trafton, Schunn, & Trickett, 2014). However, less attention has been given to differences in task comprehension within non-expert populations and how these differences may influence experimental results. In a study involving non-expert participants, Fiore and colleagues (2019) found that variations in how training is administered led to significant differences in subsequent performance on a decision-making task. This highlights the importance of understanding how varying forms of training may influence task comprehension and potentially distort the apparent effectiveness of uncertainty visualizations. However, a related study found that regardless of variations in training administration, participants who demonstrated relatively high training comprehension outperformed those who demonstrated relatively low training comprehension (Song et al., 2018). In addition to considering differences in decision making and task familiarity between experts and non-experts, we also propose that equal consideration should be given when it comes to differences *within* non-expert participant samples.

Because task comprehension may significantly influence performance, one way to mitigate this issue is to provide some means of supporting task comprehension with interactive task environments. Recent empirical research has found that modifying a computer-based problem solving task to allow for external cognition through virtual interaction with task components improved the test's ability to predict academic performance (Bocanegra, Poletiek, Ftitache, & Clark, 2019).

Given the challenges involved in this area of research, we turn to extended cognition as a theoretical framework to support the design and evaluation of uncertainty visualizations in decision support systems.

Extended Cognition to Support Problem Solving

Recently, extended and enactive cognition theory has been proposed as a theoretical foundation for understanding and improving uncertainty visualizations (Newton, Fiore, & LaViola, 2017). The concept of extended cognition (Clark & Chalmers, 1998) suggests that cognition be viewed as distributed across an operator and the system being used. For example, complex operational environments rely on decision-aiding to support activities like course of action (COA) selection. Through this lens, cognition is not simply in the decision maker, but is distributed across them and the systems used to enable that decision making. In particular, cognitive artifacts, whether low or high tech, are one piece of a hybrid human-machine system. In this way, cognitive artifacts can be seen as a form of external cognition (Fiore & Wiltshire, 2016).

Interaction with representations adds another dimension to the manipulation of visualizations to support COA selection. Interaction design, in the context of representations, takes the form of determining what actions a user should be able to take to engage with the represented information. Specifically, task environments that allow interaction with cognitive artifacts can be used to improve task understanding while reducing cognitive load through offloading. As such, our work is also theoretically driven by the notion of enactive cognition. Here, there is a necessary distinction to be made between cognition merely as the processing of information, and cognition in action (Goodwin, Wiltshire, & Fiore, 2015). In the enactive account, the environment is argued to be perceived in terms of the action possibilities that are available to an organism (Gallagher & Varga, 2014). Associated with this is the idea that embodied, active engagement with the visualization is more likely to lead to better comprehension of information when compared to lower degrees of embodied interaction (Johnson-Glenberg, Birchfield, Tolentino, & Koziupa, 2014; Gallagher & Lindgren, 2015). This, in part, drives the distinction between passive and interactive representations.

Related to human factors research, extended and enactive views of cognition align with the principles of ecological interface design (Vicente, 2002) in that its proponents emphasize the role of interaction in supporting decision making. We similarly suggest that an interface should afford interaction with visualizations to support understanding of a situation when dealing with uncertainty. But we suggest it is the actual interaction itself that constitutes cognition; that is, it is not simply affording interaction, it is the interaction that fosters understanding of uncertainty. In cognitive science theorizing, Hutchins's (1995) framework for distributed cognition similarly suggests a more complex and integrated view of cognitive processing as spread over individuals, teams, and the environment in which they operate.

Putting this in context, static representations of uncertainty are limited in that although they may depict possible outcomes, they cannot convey temporal ambiguity as effectively as dynamic and interactive representations. In contrast, dynamic representations, such as animations, allow users to view possible outcomes over time. Furthermore, interactive dynamic representations could allow users greater control through active engagement with the visualization. Here, it is the interaction with a system that creates the context for cognition to be distributed across the operator and the system, forming an externalized, extended human-machine cognitive system.

Summary

With the above as theoretical foundation, in this study, we aim to examine the extent to which training comprehension alters decision making, and to explore the effectiveness of different types of uncertainty representations, along with varying levels of interaction with dynamic uncertainty visualizations. Additionally, we propose that complex decision making can be supported by systems that afford external cognition through interactions within the task environment. In doing so, we seek to demonstrate how extended cognition theory can be applied to the design of interactive visualization interfaces to support decision making.

METHODS

The present study has two overarching goals. First, this study aims to replicate recent findings emphasizing the importance of training in uncertainty visualization research (e.g., Fiore et al., 2019; Song et al., 2018). In doing so, we contribute to this growing area of research on how to improve the design of empirical evaluations of uncertainty visualizations. Second, this study extends previous work by introducing a more complex task environment to examine the role of external cognition in decision making using an interactive interface with varying visualizations of uncertainty.

Participants

We recruited 240 participants (45.4% female and 54.6% male, mean age = 35.01 years) through Amazon's Mechanical Turk (AMT). Participants were required to be at least 21 years old due to alcohol-themed content, and were compensated $2 USD, with a bonus of $1 for participants who scored in the top ten percent during the decision-making task.

Experimental Design

We used a between-subjects design to investigate the effects of task comprehension, visualization type, and external cognition on decision making in simulated scenarios.

Independent Variables

Training comprehension. To assess how well participants retained information from the training, we administered a Knowledge Acquisition assessment prior to beginning the decision-making task (see Fiore et al., 2019 for details). This assessment consisted of a Recognition and Declarative Knowledge section to test participants' ability to identify images from the training and recall key facts from the training, respectively. Both sections had ten multiple-choice questions for a total of twenty Knowledge Acquisition assessment questions.

Visualization representation. To assess the effect of different representations of uncertainty, we used two versions of the uncertainty visualization (Figure 1). First, the Cone visualization depicted potential travel paths of a given object with a solid, translucent area. In contrast, the Line visualization depicted the same information using multiple solid lines to form a spaghetti plot. These lines indicated individual paths an object could follow. Although training for both versions specified that an object could end up anywhere in its uncertainty area, the lines may provide a more salient reminder that there are multiple possible outcomes. Both visualizations were otherwise identical.

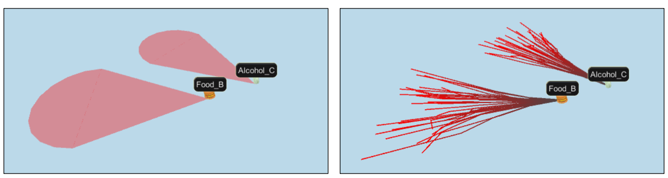

Figure 1. Uncertainty represented by cones (left) and lines (right).

Visualization interactivity. To explore how extended and enactive cognition plays a role in complex decision making, we manipulated degree of interactivity with the visualization interface. First, the Static visualization depicted uncertainty with still images. Next, the Passive Dynamic visualization allowed participants to animate potential outcomes by pressing a play button. The animation repeated on a loop, showing a different potential outcome each time, but offered no control aside from playing and stopping. Third, the Interactive Dynamic visualization depicted different potential outcomes, but allowed greater control via a slider that participants could use to move forward and backward in time.

External cognition. In addition to the dynamic visualizations described above, we measured participants' external cognition independent of their assigned condition by examining their use of cognitive artifacts. All versions of the interface allowed participants to access cognitive artifacts in the form of memory aids, or "cheat sheets" displaying information about the game components. Each memory aid provided information about a different concept (e.g., resources, assets; see Figure 2), and participants could access any of these throughout the vignettes. In line with extended cognition theory, this artifact allows decision makers to offload memory into the environment (Clark & Chalmers, 1998), thus freeing up cognitive resources to support more complex decision making. To assess the effect of externalizing and offloading information on task performance, we measured the mean frequency of cognitive artifact engagement for each vignette, as well as the mean amount of time spent engaging with the cognitive artifacts for each vignette.

Your Vehicles

	Crew	Speed (knots)	Range (km)	Capacity (kg)	Weight (kg)
Yacht	6	40	500	600	3500
RIB	2	90	250	100	400
Cruiser	3	75	300	250	800
Jet Ski	1	80	70	30	200

Supplies

Supply Type	Supply Name	Value	Weight	Value/Weight
Alcohol	Alcohol_A	3000	300	10
	Alcohol_B	2500	250	10
	Alcohol_C	250	30	8.33
	Alcohol_D	500	50	10
Food	Food_A	2500	500	5
	Food_B	250	25	10
Miscellaneous Supplies	Miscellaneous_A	1500	100	15
	Miscellaenous_B	50	5	10

Figure 2. Examples of cognitive artifacts displaying information about resources (left) and assets (right).

Decision-Making Task

To gauge knowledge transfer and examine the effect of our independent variables on decision making, we administered a Knowledge Application assessment in which participants were required to apply concepts from the training to new settings. This assessment took the form of a COA selection task designed to be an analogue to a Naval Intelligence Unit drug interdiction operation. The Knowledge Application assessment was presented as a game comprising twenty scenarios, or vignettes, involving varying combinations of party supplies lost at sea, available vehicles to retrieve them, and hazards. Participants were required to select the course of action that would maximize supplies as efficiently as possible. Each vignette was accompanied by a multiple-choice question asking participants to select the optimal answer out of four choices, while considering that multiple answers may technically be correct. Of the response options, one answer was optimal, two were technically allowed but suboptimal, and one was completely incorrect in that it violated the rules of the game and would result in loss of resources.

Dependent Variables

We used objective and subjective measures of performance during this task, as well as combinatory measures, as described below.

Performance Accuracy. In order to gauge how well participants could apply the knowledge gained during training to a complex decision-making task, we calculated the mean percent of optimal decisions for the vignettes.

Subjective Workload. Immediately following each assessment item, participants were asked to rate how difficult it was for them to answer the question using a 7-point Likert-type scale with "very easy" and "very difficult" as anchors.

Cognitive Efficiency. By combining both objective (i.e., accuracy) and subjective (i.e., workload) measures, we derived scores for participants' cognitive efficiency (see Fiore, Scielzo, Jentsch, & Howard, 2006), we calculated the relationship between standardized accuracy and standardized workload scores, such that a positive score indicates higher accuracy relative to workload, and a negative score indicates lower accuracy relative to workload.

Metacognitive Bias. We calculated an additional score combining objective and subjective performance measures by assessing the relationship between participants' predicted accuracy and their actual accuracy. Previous work has investigated how task comprehension may influence assessments of one's own performance on a complex task (Cuevas, Fiore, Bowers, & Salas, 2004). Participants were asked to predict how many questions they would answer correctly based on how well they understood the training. Participants provided both a prediction immediately before the assessment, as well as a postdiction of how well they thought they had performed immediately following the assessment. Metacognitive bias was measured by subtracting actual performance from predicted performance. Thus, a positive score indicates overestimation of performance, while a negative score indicates underestimation.

Hypotheses

H1: Training Comprehension. Participants with higher scores on the Knowledge Acquisition assessment will perform better on the Knowledge Application assessment (simulation vignettes). Specifically, higher training comprehension will lead to higher accuracy, lower workload, lower metacognitive bias, and higher cognitive efficiency.

H2: Visualization representation and interactivity. Both the type of visualization and the level of interaction afforded by the visualizations will have an effect on vignette performance. Specifically, Line visualizations will lead to improved performance relative to Cone visualizations, and Interactive Dynamic visualizations will lead to improved performance relative to the Passive Dynamic and Static visualizations as defined by higher accuracy, lower workload, lower metacognitive bias, and higher cognitive efficiency.

H3: External Cognition. Participants who demonstrate greater external cognition through higher engagement with cognitive artifacts in the vignette interface will exhibit better performance on the Knowledge Application assessment, as defined by higher accuracy, lower workload, lower metacognitive bias, and higher cognitive efficiency.

Procedure

Participants were directed from AMT to our study hosted on Qualtrics, where they were randomly assigned to one of six visualization conditions that differed in type of uncertainty representation and interactivity (2 x 3 factorial). All participants completed a training on the decision-making scenario and its components (for a more detailed description of how the training was developed, see Fiore et al., 2019). Participants in all conditions completed the same training, which consisted of written instructions, images, and diagrams describing game components and objectives. Following this, participants completed a task interface tutorial. Participants in the Static condition received instructions on how to use the interface and answer the vignette questions. Participants in the Passive Dynamic condition also received these instructions, as well as instructions on how to play animations for the vignettes. Finally, participants in the Interactive Dynamic condition received the same instructions as the Static condition, as well as instructions on how to manipulate the slider to move forward and backward in time for the vignettes.

RESULTS

We conducted regression analyses using R (R Core Team, 2018) to examine the effect of our independent variables on performance on the simulation vignettes. First, a multivariate multiple linear regression analysis showed our IVs significantly predicted accuracy, workload, and post-metacognitive bias. There was no significant effect on pre-metacognitive bias. Additionally, for our combinatory metric of relative accuracy to workload, a multiple linear regression showed our IVs significantly predicted cognitive efficiency. Table 1 depicts a summary of significant predictors for each dependent variable.

Table 1. Summary of significant predictors of vignette performance.

	Training Comprehension		External Cognition Frequency			
DV	*B (SE)*	β	*B (SE)*	β	R^2	*F*
Accuracy	.11 (.03)	.17**	.04 (.02)	.31*	.13	5.70***
Workload	-1.69 (.44)	-.25***	-.33 (.22)	-.23	.10	4.19***
Metacognitive Bias - Pre	-.12 (.08)	-.10	.02 (.04)	.08	.02	.94
Metacognitive Bias - Post	-.26 (.08)	-.23***	-.02 (.04)	-.08	.05	2.15*
Cognitive Efficiency	1.37 (.32)	.27***	.37 (.16)	.34*	.16	7.53***

*Sig: * p < .05, ** p < .01, *** p < .001*

Training comprehension. In support of H1, higher training comprehension significantly predicted higher vignette accuracy (β = .17), as well as lower workload (β = -.25). While it did not have a significant effect on prediction metacognitive bias prior to the vignettes, higher training comprehension significantly reduced postdiction metacognitive bias following the vignettes (β = -.23). Additionally, for cognitive efficiency, which was derived from accuracy and workload scores, higher training comprehension significantly predicted higher cognitive efficiency (β = .27).

Visualization type. H2 was not supported. There was no significant effect of visualization representation or interactivity on vignette performance.

External cognition. In partial support of H3, higher frequency of cognitive artifact use significantly predicted higher accuracy (β = .31), and while it did not have a significant effect on workload, higher frequency of cognitive artifact use significantly predicted higher cognitive efficiency, indicating higher relative accuracy to workload (β = .34). There was no significant effect of time spent engaging with cognitive artifacts on accuracy, workload, or cognitive efficiency, nor was there a significant effect of either frequency or time of cognitive artifact use on metacognitive bias.

DISCUSSION

In line with previous work, participants who demonstrated higher training comprehension also demonstrated superior knowledge transfer during decision-making task. Critically, regardless of differences in the task environment, such as visualization representation and interface interactivity, higher scores on the Knowledge Acquisition assessment significantly predicted performance on the Knowledge Application assessment, as indicated by higher accuracy, lower workload, lower metacognitive bias following completion of the assessment, as well as higher cognitive efficiency. This finding further illustrates how evaluations of uncertainty visualizations must consider the effect of learning on performance regardless of visualization manipulations, particularly when participants lack prior knowledge or experience related to the task.

While level of interactivity of the visualizations did not have a significant effect on performance, it is worth noting that the majority of participants in the dynamic visualization conditions did not engage with the interactive features at all. This could mean such participants treated the visualizations as if they were static, thus limiting our ability to make conclusions about the dynamic visualizations themselves. Future work can explore differences in participants' engagement with dynamic visualizations, and potential benefits and trade-offs of using such features.

Interestingly, although the present study did not assess performance differences between experts and non-experts, our findings indicate *task-specific* expertise differences among participants, as determined by training comprehension, may produce fundamental differences in performance. Importantly, the present study allowed us to investigate the role of externalizing cognitive processes during complex tasks. Our findings suggest that participants who offloaded via the use of cognitive artifacts exhibited improved decision-making accuracy. Additionally, external cognition increased cognitive efficiency, suggesting that not only was accuracy improved, but that it was improved without increasing workload. This finding supports the argument that externalizing cognition can reduce working memory demands, thereby supporting complex cognition by increasing available cognitive resources.

While we found partial support for our hypothesis that external cognition improves performance, recent work suggests alternating periods of external and internal cognition, indicated by distributions of actions relative to inactive periods, are more predictive of superior problem solving than net time or frequency of actions (Bocanegra et al., 2019). Because we only measured overall time and frequency of external cognition, our analyses did not consider temporal variations in externalizing actions, which may have accounted for the relatively modest effect of external cognition we found. Thus, future work is needed to investigate whether patterns of actions indicating shifts between external and internal cognition are a better predictor of performance than overall external cognition.

Overall, our findings suggest that participants who demonstrate greater mastery of the task training performed better during the experimental task with relatively lower workload and made more accurate judgments of their performance following the task. Additionally, our findings suggest the use of computer-generated cognitive artifacts may also improve performance, and critically, do so while maintaining high cognitive efficiency.

CONCLUSION

The challenges of decision making under uncertain conditions, along with those of designing effective decision support systems, highlight a need for continued research in this area. This study supports and builds upon previous research investigating factors that may alter both the efficacy of uncertainty visualizations, as well as empirical assessments of these visualizations, and offers support for further examination of extended cognition theory as a theoretical framework for improving the design of technologies to support complex cognition.

ACKNOWLEDGEMENT

This work was supported by the Office of Naval Research grant N00014-15-1-2708, under the Command Decision Making program. The views and opinions contained in this article are the authors and should not be construed as official or as reflecting the views of the University of Central Florida or the Office of Naval Research.

REFERENCES

Bocanegra, B. R., Poletiek, F. H., Ftitache, B., & Clark, A. (2019). Intelligent problem-solvers externalize cognitive operations. *Nature Human Behaviour*, *3*(2), 136–142. https://doi.org/10.1038/s41562-018-0509-y

Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, *58*(1), 7–19. https://doi.org/10.1111/1467-8284.00096

Cuevas, H. M., Fiore, S. M., Bowers, C. A., & Salas, E. (2004). Fostering constructive cognitive and metacognitive activity in computer-based complex task training environments. *Computers in Human Behavior*, *20*(2), 225–241. https://doi.org/10.1016/j.chb.2003.10.016

Fiore, S. M., Scielzo, S., Jentsch, F., & Howard, M. L. (2006). Understanding performance and cognitive efficiency when training for x-ray security screening. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *50*(25), 2610–2614. https://doi.org/10.1177/154193120605002501

Fiore, S. M., Song, J., Newton, O. B., Pittman, C., Warta, S. F., & LaViola, J. J. (2019). Determining the effect of training on uncertainty visualization evaluations. In T. Z. Ahram & C. Falcão (Eds.), *Advances in Usability, User Experience and Assistive Technology* (pp. 141–152). Springer International Publishing.

Fiore, S. M., & Wiltshire, T. J. (2016). Technology as teammate: Examining the role of external cognition in support of team cognitive processes. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01531

Gallagher, S., & Varga, S. (2014). Social constraints on the direct perception of emotions and intentions. *Topoi, 33*(1), 185–199.

Gallagher, S., & Lindgren, R. (2015). Enactive metaphors: Learning through full-body engagement. *Educational Psychology Review, 27*(3), 391–404.

Goodwin, M. S., Wiltshire, T., & Fiore, S. M. (2015). Applying research in the cognitive sciences to the design and delivery of instruction in virtual reality learning environments. In R. Shumaker & S. Lackey (Eds.), *Virtual, Augmented and Mixed Reality* (Vol. 9179, pp. 280–291). Springer.

Hutchins, E. (1995). How a cockpit remembers its speeds. *Cognitive Science, 19*(3), 265–288.

Johnson-Glenberg, M. C., Birchfield, D. A., Tolentino, L., & Koziupa, T. (2014). Collaborative embodied learning in mixed reality motion-capture environments: two science studies. *Journal of Educational Psychology, 106*(1), 86–104.

Kinkeldey, C., MacEachren, A. M., Riveiro, M., & Schiewe, J. (2017). Evaluating the effect of visually represented geodata uncertainty on decision-making: Systematic review, lessons learned, and recommendations. *Cartography and Geographic Information Science*, *44*(1), 1–21. https://doi.org/10.1080/15230406.2015.1089792

Kirschenbaum, S. S., Trafton, J. G., Schunn, C. D., & Trickett, S. B. (2014). Visualizing uncertainty: The impact on performance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *56*(3), 509–520. https://doi.org/10.1177/0018720813498093

Newton, O. B., Fiore, S. M., & LaViola, J. J. (2017). An external cognition framework for visualizing uncertainty in support of situation awareness. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *61*(1), 1198–1202. https://doi.org/10.1177/1541931213601782

R Core Team (2018). R: A language and environment for statistical computing. Vienna, Austria: R foundation for Statistical Computing. https://www.R-project.org

Song, J., Newton, O. B., Fiore, S. M., Coad, J., Clark, J., Pittman, C., & LaViola, J. J. (2018). Examining the impact of training and feedback on visualization-supported decision making under uncertainty. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *62*(1), 1449–1453. https://doi.org/10.1177/1541931218621329

Vicente, K. J. (2002). Ecological interface design: Progress and challenges. *Human Factors*, *44*(1), 62–78. https://doi.org/10.1518/0018720024494829