International Journal of Recent Innovations in Academic Research
This work is licensed under a Creative Commons Attribution 4.0 International License [CC BY 4.0]
E-ISSN: 2635-3040; P-ISSN: 2659-1561
Homepage: https://www.ijriar.com/
Volume-10, Issue-3, July-September-2026: 82-91

Research Article

Quantifying the Effect of Misleading Visualizations on Human Judgment with an Explainable Machine-Learning Analysis of Visual Deception and Memorability for Space Interface Design

Yujeong Yi

Independent Researcher, US International School, Seoul, South Korea
Email: yuey0319@gmail.com

Article History

Received: July 29, 2026
Accepted: August 20, 2026
Published: August 26, 2026

Abstract

The same chart that contains identical information but drawn in two different ways can lead one person to the correct conclusion and another to the wrong one. In high-risk environments such as spacecraft and mission-control systems, that difference is more than a matter of visual design. It can affect operational decisions and, ultimately, mission safety. This study investigates when information visualizations lead to judgment errors, where viewers allocate visual attention, what they remember, and whether these outcomes can be predicted from visual features. Two publicly available human datasets, CALVI and MASSVIS, were analyzed. The CALVI analysis showed that misleading visualizations reduced judgment accuracy from 79.7% to 39.4%, with accuracy varying from 3.6% to 70.1% across different types of misleading designs. The MASSVIS analysis found that viewers directed their first fixation to the data region in 68.9% of cases and that visualizations containing human-recognizable objects achieved much higher memorability than those without. Explainable machine learning models were used to predict deception and memorability, and SHAP identified the visual features that contributed most to each prediction. This study establishes quantitative evidence on how visualization design influences human interpretation and informs the design and evaluation of dependable interfaces for high-risk space environments.

Keywords: Information Visualization, Misleading Visualizations, Eye-Tracking, Memorability, Explainable Machine Learning, Human Factors, Space Interfaces.

1. Introduction

Data visualizations play an important role in high-risk environments such as spacecraft, mission control centers, and aircraft cockpits. Operators rely on visual displays to monitor system status, detect problems, and make rapid decisions. In these settings, even a small misunderstanding can lead to serious consequences. Space missions are especially demanding because operators must work under time pressure, uncertainty, and limited opportunities to recover from mistakes [1]. For this reason, the design of a visualization influences human performance heavily. If important information is difficult to find or is displayed in a misleading way, users are more likely to make incorrect decisions. NASA refers to these interface-related mistakes as design-induced errors and emphasizes that flight displays should be easy to interpret, minimize cognitive workload, and support situation awareness so that crews can respond to hazards quickly [2].

Importance of visual clarity is particularly vital in spaceflight. The NASA Space Shuttle Cockpit Avionics Upgrade redesigned fault-management displays after researchers found that cluttered interfaces made system failures harder to diagnose. Simulator studies showed that the redesigned displays improved situation awareness, reduced workload, and increased task performance [3]. Other studies also found that color helps operators identify urgent information more efficiently when it is used appropriately [4]. As space missions extend beyond low Earth orbit, communication delays and blackouts will require astronauts to make decisions

without support from ground control. Their survival will be more dependent on on-board electronic devices. NASA identifies human-computer interaction as an important operational risk for future space missions [5]. In such conditions, a misleading visualization is a potential threat to the mission and the crew.

These operational challenges and risks motivate three questions. First, what kinds of visualizations are most likely to mislead people? Second, where do people look when reading a visualization, and which visual features make information easier to remember? Third, can visual clarity be predicted from measurable characteristics of a visualization before the design is deployed? Previous studies of visual cognition show that people rely on rapid, intuitive processing, which can increase susceptibility to misleading visual encodings [6]. Also, visual characteristics influence both eye movements and memorability [7, 8]. Furthermore, several public datasets have made it possible to study human responses to visualizations using behavioral and eye-tracking data [9, 10].

This study uses similar methods to those used to evaluate spaceflight interfaces. NASA's Intelligent Spacecraft Interface Systems laboratory recorded astronaut eye-tracking with behavioral measures such as error rates and workload to understand how crews search for information and make decisions during simulated spaceflight [11]. These measurements were also used to develop models that predict human eye movements. Since then, eye-tracking has become a common tool for evaluating attention, workload, and situation awareness in aviation and astronaut training [12, 13].

There are two datasets used in this study that follow a similar objective. The CALVI dataset measures how accurately people interpret misleading visualizations [9], and the MASSVIS dataset contains eye-tracking with memorability data from real-world visualizations [7, 8]. These datasets contribute to studying judgment accuracy, visual attention, and memory using large-scale human data. They also provide an efficient way to investigate design problems that are difficult to study in spaceflight environments. Using CALVI and MASSVIS, we first examine how different misleading designs affect judgment accuracy. We then analyze how visual attention is distributed across a visualization and which visual features are associated with memorability. Finally, we develop explainable machine learning models to predict judgment errors and memorability and use SHAP to identify the visual features that contribute most to each outcome.

2. Related Work

2.1. Visualization and Decision Making

Studies have shown that the way information is visualized affects how people interpret data. Padilla et al. [6] proposed a cognitive framework based on dual-process theory, which suggests that people often rely on fast, intuitive processing before engaging in slower, more accurate reasoning. As a result, visual features that immediately attract attention can strongly influence judgment, especially under time pressure. For instance, operators in spacecraft or mission control need to make rapid decisions using visual displays. If a visualization emphasizes the wrong information or makes important information difficult to notice, users may reach an incorrect conclusion even with accurate data.

2.2. Misleading Visualizations and Visualization Literacy

Visualization literacy has focused on whether people can correctly interpret well-designed charts. Ge et al. [9] pointed out that this approach overlooks the ability to recognize a misleading visualization. They developed CALVI, a visualization literacy assessment that contains eleven categories of misleading designs, such as truncated axes, inverted scales, and other forms of visual distortion in multiple chart types. The dataset contains misleading visualizations and a large collection of human responses. Their results showed that susceptibility to misleading visualizations can be measured more precisely and varies across different types of misleading designs. In the context of space systems, similar design choices could unintentionally appear in spacecraft telemetry displays or mission-control interfaces. Understanding which types of misleading visualizations produce the highest error rates can help identify design patterns that require additional revision during interface development.

2.3. Attention, Memory, and Visual Design

Another research investigates how visual design influences attention and memory. Borkin et al. [7] found that some visual attributes are remembered more consistently than others, with human-recognizable objects showing one of the strongest effects on memorability. In a follow-up study, they combined eye-tracking with memorability data and showed that viewing patterns are closely related to what people later recognize and recall [8]. The MASSVIS dataset from these studies includes eye-tracking data, memorability scores, and annotations describing the visual properties of real-world visualizations. The dataset makes it possible to

examine both where people pay their attention and which visual attributes are associated with stronger memory. These questions are also relevant to spaceflight interfaces because operators must quickly identify important information and recall it while completing complex tasks.

2.4. Eye Tracking and Predictive Modelling of Operator Behaviour

Eye-tracking is used to evaluate attention and situation awareness in safety-critical environments. Studies in aviation have shown that experienced pilots scan displays differently from novices, and these differences are associated with better situation awareness and task performance [12]. Similarly, NASA's Intelligent Spacecraft Interface Systems laboratory combined astronaut eye-tracking with behavioral measures such as error rates and workload during simulated missions. These data were then used to develop models that predict eye movements [11]. Reviews of astronaut training have also identified eye-tracking as an effective way to estimate attention and mental workload during tasks such as spacecraft docking and teleoperation [13]. Like previous research, we combine eye-tracking and behavioral data with machine learning. However, rather than focusing only on prediction accuracy, we use SHAP to explain which visual features contribute most to judgment errors and memorability.

2.5. Human-Centered Datasets for Visualization Understanding

Recent public datasets have made it easier to study how people interpret visualizations. Verma et al. [10] introduced CHART-6, a benchmark that combines several visualization datasets with corresponding human responses. Using these methods, we analyze how people interpret visualizations using observed behavior. With the MASSVIS dataset [7, 8], these datasets provide information about judgment, attention, and memory. We examined multiple aspects of human visualization understanding.

2.6. Human Factors in Space Interfaces

Research in aerospace human factors has consistently shown that interface design affects operator's performance. NASA's human integration standards identify design-induced error, cognitive workload, and situation awareness as important considerations when developing crew interfaces [2]. The Cockpit Avionics Upgrade program also showed that redesigning cluttered displays improved situation awareness, reduced workload, and increased task performance, with color coding playing an important role [3, 4]. As future missions require greater crew autonomy, NASA also recognizes human-computer interaction as an important operational risk [5].

3. Methodology

3.1. Overview of Studies

This paper consists of three complementary studies that investigate different aspects of how people interpret information visualizations. Study 1 analyzes the CALVI dataset to measure how misleading chart designs affect judgment accuracy. Study 2 analyzes the MASSVIS dataset to examine where viewers direct their attention the most and which visual attributes improve memorability.

Study 3 builds explainable machine learning models to predict deception and memorability from visual features and uses SHAP to identify the features that drive each outcome. Overall, these three studies identifies when visualizations cause errors, explain how people process visual information, and finally predict these outcomes using explainable models.

3.2. Datasets

This study uses two publicly available human datasets. The first dataset CALVI focuses on judgment errors caused by misleading visualizations, and the second dataset MASSVIS focuses on visual attention and memory. They are analyzed independently because the datasets have different stimuli, tasks, and participants.

3.2.1. CALVI (Deception)

The first dataset is the Critical Thinking Assessment for Literacy in Visualizations (CALVI) [9], from the CHART-6 benchmark [10]. CALVI measures how well people identify misleading visualizations. Each item consists of a chart and a corresponding question. Some charts are intentionally misleading, and others are correctly designed. Misleading charts include eleven types of visual distortions, such as inappropriate axis ranges and inverted axes. Each response records whether the participant answered correctly. From the CHART-6 human-response dataset, we selected all records with the CALVI test type and Human agent type. The final dataset contains 13,926 responses with 60 visualization items. Among these, 7,485 responses correspond to normal visualizations and 6,441 correspond to misleading visualizations. The 45 misleading items are also labeled by misleader type (11 categories) and chart type (9 categories).

3.2.2. MASSVIS (Attention and Memory)

The second dataset is the Massachusetts Visualization (MASSVIS) dataset [7, 8], which combines real-world visualizations with memorability scores, eye-tracking data, and visual attribute annotations. We analyzed the 393 visualizations for which all of these data are available. Each visualization contains annotations about its visual properties, including data-ink ratio, number of distinct colors, visual density, black-and-white status, the presence of human-recognizable objects or human figures, and data and message redundancy.

Memorability is measured using the prolonged hit rate, which is the percentage of participants who correctly recognized a visualization after a delay. The eye-tracking data record the location and duration of every fixation while participants viewed each visualization. Additional annotation files explain the locations of interface elements such as the title, data region, axes, legend, and other components. These annotations allow each fixation to be assigned to a specific region of interest for further analysis.

3.3. Study 1: Deception Analysis (CALVI)

We investigated how misleading visualizations affect judgment accuracy. Each response was classified as either a normal or misleading visualization. To examine this, we compared accuracy between the two groups using the Mann-Whitney U test. In the second analysis, we matched each misleading response with its corresponding misleader type. We calculated the accuracy for each misleader category and used a chi-square test of independence to elucidate whether judgment accuracy differed across misleader types. We also calculated accuracy for each individual visualization to identify the designs that produced the highest error rates.

3.4. Study 2: Attention and Memory Analysis (MASSVIS)

We examined how visual design influences memory and visual attention. We first tested the relationship between each visual attribute and memorability. Binary attributes, including the presence of a human-recognizable object and data or message redundancy, were compared using the Mann-Whitney U test. Continuous and ordinal attributes, which includes the number of distinct colors, data-ink ratio, and visual density, were analyzed using Spearman's rank correlation.

For the attention analysis, each fixation was assigned to a region of interest according to its coordinates. The labeled interface elements were grouped into five regions including the title, data region, axis, legend, and other elements. We identified both the first fixation and all subsequent fixations for every participant and visualization. From 393 visualizations, we calculated the proportion of first fixations, the proportion of total fixations, and the mean fixation duration for each region. These measures were used to examine where participants looked first and how visual attention was distributed throughout each visualization.

3.5. Study 3: Predictive Modelling and Explanation

In the third study, we assessed whether visual features can predict memorability and judgment accuracy. We developed two separate prediction tasks, one using the MASSVIS dataset and the other using the CALVI dataset. For the MASSVIS dataset, each visualization was represented by eight visual attributes. Binary variables were encoded as 0 or 1, and memorability was classified as high or low using the median prolonged hit rate (84.2%) as the threshold. For the CALVI dataset, each response was represented by the misleader type, chart type, and task category. These variables were one-hot encoded and had 26 input features.

The prediction target was whether the participant answered correctly or was misled by the visualization. For both datasets, we trained a LightGBM classifier and compared its performance with a Logistic Regression baseline. Model performance was evaluated using five-fold cross-validation with accuracy, F1 score, and ROC-AUC. To explain the model predictions, we used SHAP (SHapley Additive exPlanations).

SHAP estimates the contribution of each feature to the model's predictions and ranks features according to their overall importance. This analysis identifies which visual characteristics contribute most to memorability and judgment errors. Table 1 summarizes the datasets, analyses, and methods used in each study.

Table 1. Summary of datasets and analytical methods across the three studies.

Study	Dataset	Unit of analysis	Primary methods
1 (Deception)	CALVI [9, 10]	13,926 responses / 60 items	Mann-Whitney U; chi-square
2 (Memory)	MASSVIS [7, 8]	393 visualizations	Mann-Whitney U; Spearman
2 (Attention)	MASSVIS [7, 8]	Fixations over 393 vis.	Point-in-polygon AOI mapping
3 (Prediction)	Both	Responses / visualizations	LightGBM vs logistic; SHAP

4. Results

The analyses use behavioral responses from 497 participants in the CALVI dataset and eye-tracking data collected from 33 participants across 393 visualizations in the MASSVIS dataset.

4.1. Study 1: Misleading Visualizations and Judgment Errors (CALVI)

4.1.1. Misleading Design Substantially Reduces Judgment Accuracy

Human accuracy was compared between normal and misleading visualizations. Participants achieved a mean accuracy of 79.7% on normal charts (n = 7,485), compared with 39.4% on misleading charts (n = 6,441), shown in Figure 1. A Mann-Whitney U test showed that this difference of 40.3 percentage points was statistically significant ($p < 0.001$). Although the underlying data were identical, changing only the visual encoding reduced the likelihood of a correct judgment by nearly half. Visual design alone can substantially influence human interpretation in situations where accurate decisions depend on rapid interpretation of graphical information.

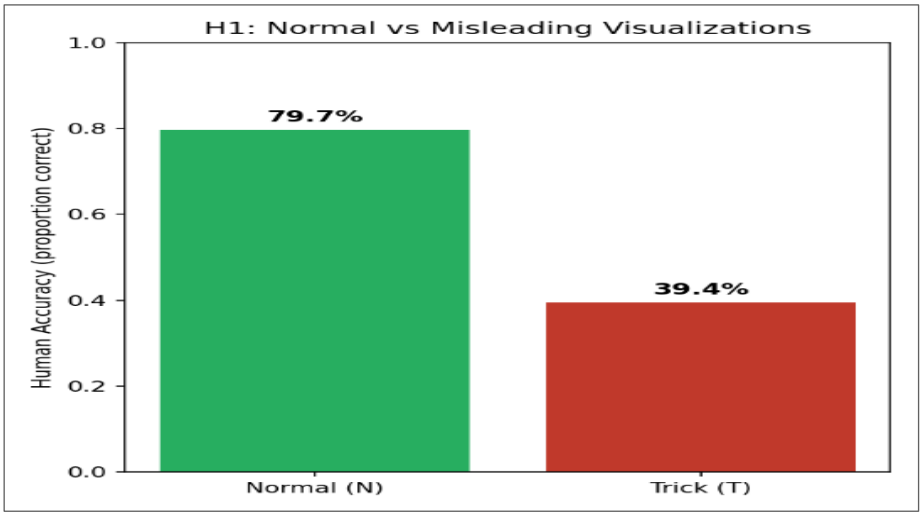

Figure 1. Human accuracy on normal versus misleading visualizations.

4.1.2. Deceptiveness Varies Substantially Across Misleader Types

The 45 misleading visualizations were grouped into 11 misleader categories, and judgment accuracy was calculated for each category. Accuracy ranged from 3.6% to 70.1%. A chi-square test confirmed that the differences among categories were statistically significant (chi-square = 1113.4, df = 10, p is approximately 10^-233). Figure 2 reports the most and least deceptive types.

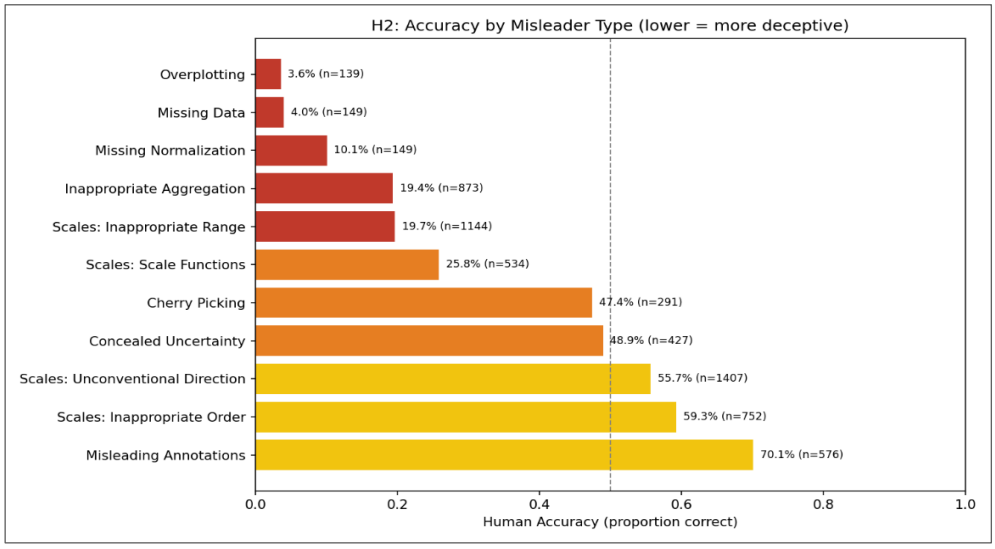

Figure 2. Human accuracy by misleader type (selected types, ordered from most to least deceptive).

The results further support the idea that some misleading designs consistently produced much lower accuracy than others. Overplotting and Missing Data were associated with the lowest performance, whereas categories involving reordering or annotation produced relatively lower reductions in accuracy. Thus, different forms of misleading visualization vary considerably in their effect on human judgment.

4.2. Study 2: Visual Attention and Memory (MASSVIS)

4.2.1. Visual Attributes Influence Memorability

The relationship between visual attributes and memorability was examined across 393 visualizations (mean memorability = 79.7%). Table 2 summarizes the results for the three binary attributes.

Table 2. Memorability by binary visual attribute.

Attribute	Present	Absent	p-value
Human recognizable object (HRO)	90.5%	73.4%	~10^-24
Message redundancy	81.8%	78.4%	0.001
Data redundancy	78.5%	80.0%	0.80 (n.s.)

According to Figure 3, visualizations containing a human-recognizable object achieved memorability scores that were 17.1 percentage points higher than those without such objects ($p < 0.001$). Message redundancy was also associated with higher memorability, although the effect was smaller. Data redundancy was not significantly associated with memorability. Among the continuous visual attributes, the number of distinct colors (Spearman's $\rho = 0.282$), data-ink ratio ($\rho = 0.190$), and visual density ($\rho = 0.170$) all showed weak positive correlations with memorability. Semantic visual elements showed a stronger association with memorability than redundant information alone.

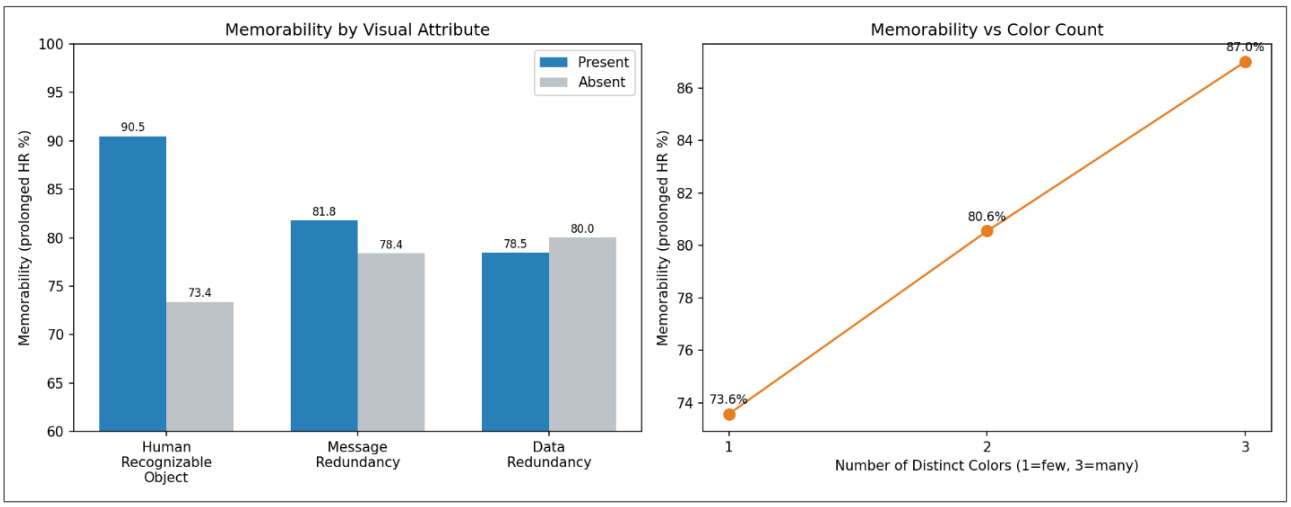

Figure 3. Memorability by visual attribute and by number of distinct colors.

4.2.2. Initial Fixations Are Directed Predominantly to the Data Region

Each fixation was assigned to one of five regions of interest including the title, data region, axis, legend, and other areas. The distribution of first fixations and total fixations was then analyzed. Figure 4 shows the proportion of first fixations and of all fixations allocated to each region.

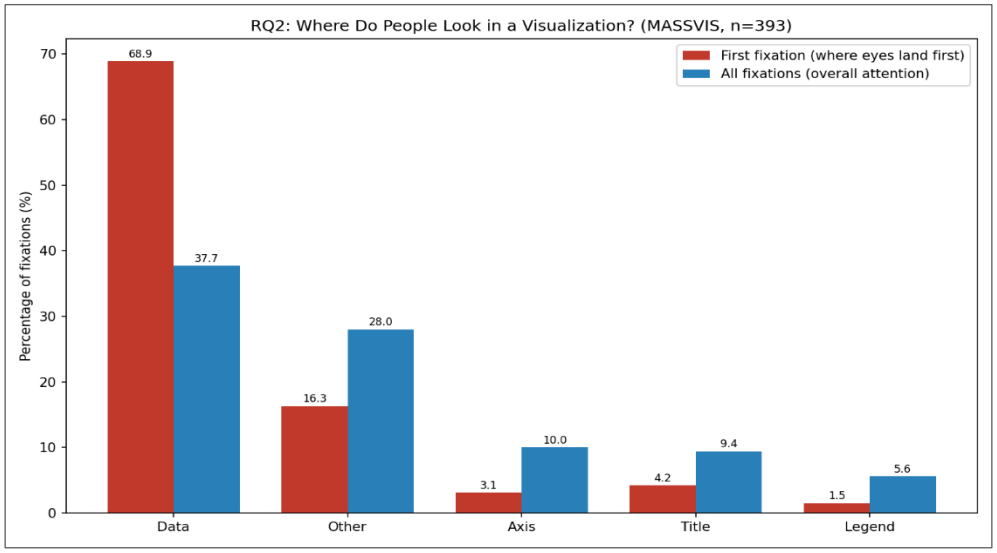

Figure 4. Distribution of first and total fixations across regions of interest.

The first fixation landed on the data region in 68.9% of observations, compared with 4.2% for the title. The data region also received the largest proportion of all fixations (37.7%). Mean fixation duration ranged from 205 ms to 231 ms across all regions, with relatively small differences. This suggests that visual attention differed mainly in where participants looked rather than in how long they looked at each location.

4.3. Study 3: Prediction and Explanation (Machine Learning and SHAP)

4.3.1. Visual Features Predict Deception and Memory, With Task-Dependent Structure

Two prediction tasks were evaluated. The first task classified each visualization in the MASSVIS dataset as having above- or below-median memorability (84.2%) using eight visual attributes. Table 3 summarizes the model performance.

Table 3. Memorability classification performance (5-fold cross-validation).

Model	Accuracy	F1	ROC-AUC
Baseline (logistic regression)	0.697	0.657	0.726
LightGBM	0.659	0.638	0.680

For memorability prediction, Logistic Regression achieved an accuracy of 0.697 and a ROC-AUC of 0.726, while LightGBM achieved an accuracy of 0.659 and a ROC-AUC of 0.680. The linear baseline performed slightly better than the gradient-boosting model. SHAP analysis identified the presence of a human-recognizable object as the most influential feature (mean absolute SHAP = 0.987), which is consistent with the statistical results reported above. Figure 5 shows that the presence of a human-recognizable object was the dominant feature for predicting memorability.

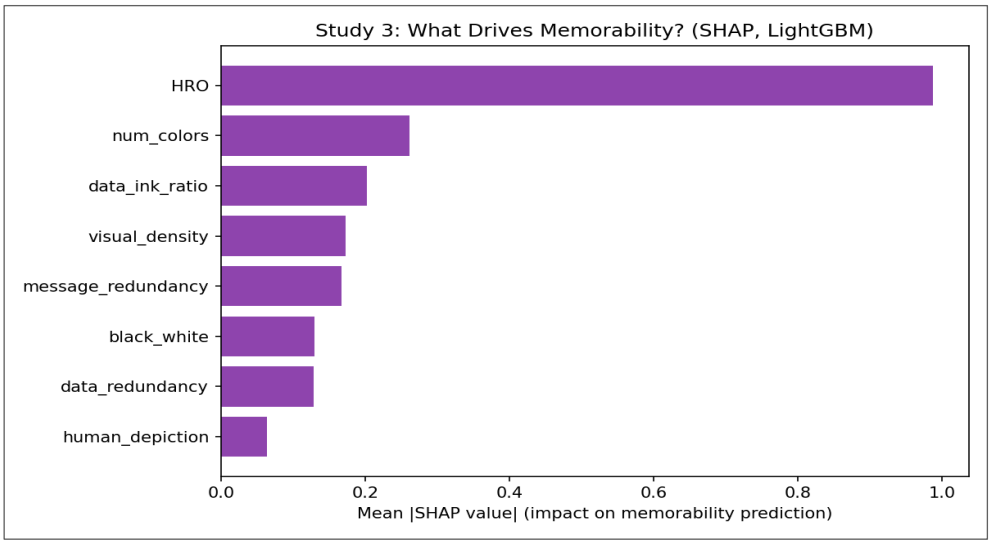

Figure 5. SHAP feature importance for memorability prediction (MASSVIS).

The second task predicted whether participants correctly interpreted or were misled by each visualization in the CALVI dataset using 26 encoded visual features.

Table 4 summarizes the model performance.

Table 4. Deception classification performance (5-fold cross-validation).

Model	Accuracy	F1	ROC-AUC
Baseline (logistic regression)	0.747	0.642	0.797
LightGBM	0.758	0.651	0.818

For deception prediction, Logistic Regression achieved an accuracy of 0.747 and a ROC-AUC of 0.797. LightGBM achieved an accuracy of 0.758 and a ROC-AUC of 0.818, outperforming the linear baseline across all evaluation metrics. SHAP analysis identified Misleading Annotations, several Scale Manipulation subtypes, and Bar Chart encodings as the most influential features. Memorability was predicted more effectively by a linear model, whereas deception prediction benefited from the nonlinear relationships captured by LightGBM. Figure 6 presents the SHAP importance values, identifying misleading annotations and scale manipulations as the strongest drivers of deception.

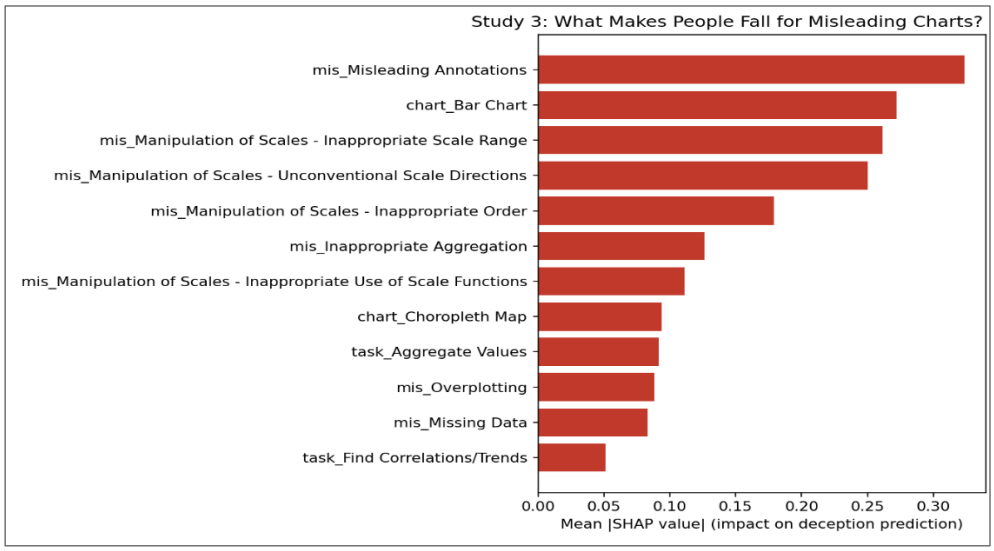

Figure 6. SHAP feature importance for deception prediction (CALVI).

5. Discussion

5.1. Interpretation of Findings

The results from Study 1 illustrate that misleading visualizations substantially reduce judgment accuracy. Accuracy decreased from 79.7% for normal visualizations to 39.4% for misleading visualizations. However, not all misleading designs had the same effect. Some misleader types produced much higher error rates than others, ranging from 3.6% accuracy for overplotting to 70.1% for misleading annotations. Overplotting, missing data, and inappropriate scale ranges were among the most deceptive designs, yet misleading annotations had a much smaller effect. Thus, certain visualization techniques are consistently more likely to produce judgment errors.

The findings from Study 2 help explain why some misleading designs are more effective than others. Participants directed their first fixation to the data region in nearly 68.9% of cases, much more often than to the title (4.2%) or other interface elements. Misleading information placed within the data region is likely to influence interpretation early in the viewing process. The memory analysis also showed that visualizations containing human-recognizable objects were remembered 17.1 percentage points more often than those without them. In contrast, data redundancy did not significantly improve memorability. As a result, meaningful visual content contributes more to long-term memory than repeating the same information.

Study 3 further corroborates these findings. Memorability was predicted reasonably well (area under the ROC curve 0.73), but the LightGBM model did not outperform the Logistic Regression baseline. This suggests that the relationship between the measured visual attributes and memorability is relatively simple. SHAP analysis also identified human-recognizable objects as the most influential feature, consistent with the statistical results from Study 2. The results for deception showed a different pattern. LightGBM achieved better performance than Logistic Regression (area under the ROC curve 0.82). Judgment errors depend on interactions between multiple visual features rather than on individual features alone. SHAP identified several misleading design types as the strongest predictors of deception, which is consistent with the statistical analysis in Study 1. The agreement between the statistical tests and the machine learning models increases confidence in the overall findings.

5.2. Design Principles for High-Risk Space Interfaces

These findings can be used as a guidance for the design and verification of spacecraft and mission-control interfaces, where interpretation errors can have serious consequences.

First, interface verification should focus on the most deceptive visualization designs. Since different misleader types vary greatly in their effect on judgment accuracy, the review process should prioritize designs that suppress or distort the underlying data, particularly overplotting, missing data, and inappropriate scale ranges. The ranking of misleader types identified in this study can serve as a useful reference during interface review. Second, safety-critical information should be placed where users naturally direct their attention. The data region receives most first fixations, and therefore, critical information should appear in this area rather than in titles or legends that may receive attention much later. Likewise, misleading elements located in the

data region are likely to have the greatest impact because they are seen immediately. Third, information that operators must remember should be supported by meaningful graphical elements. The strong memory advantage of human-recognizable objects, together with the negligible effect of data redundancy, suggests that critical information is more effectively communicated through recognizable icons or graphical markers than through repeated text or numerical values. This result is consistent with previous work showing that color and graphical coding improved situation awareness and reduced workload in redesigned Space Shuttle cockpit displays [3, 4].

Finally, explainable machine learning could support interface evaluation during system development. Because deception could be predicted from visual features and its main drivers were identified through SHAP, an explainable model could help identify high-risk display configurations before deployment while also explaining why a particular design is likely to cause errors. It is consistent with previous NASA studies that used eye-movement and performance models to evaluate spacecraft interfaces [11]. Unlike a black-box model, an explainable model provides information that can support interface improvement.

6. Conclusion

This study analyzed two publicly available human datasets to elucidate when visualizations lead to judgment errors, where viewers direct their attention, what they remember, and whether these outcomes can be predicted using machine learning. The results showed that misleading visualizations substantially reduced judgment accuracy, with the largest effects produced by visualization designs that distort or hide the underlying data. The eye-tracking analysis showed that viewers directed their attention primarily to the data region, and the memorability analysis showed that meaningful graphical elements improved long-term memory more effectively than data redundancy.

The machine learning analysis further demonstrated that deception and memorability can be predicted from visual features, and SHAP identified the visual characteristics that contributed most to these predictions. These findings can be utilized for designing high-risk interfaces. Interface verification should focus on the most deceptive visualization designs; critical information should be placed where users naturally direct their attention; information that must be remembered should be supported by meaningful graphical elements; explainable machine learning can assist and predict in identifying potentially risky interface designs during development.

Declarations

Acknowledgments: The author gratefully acknowledges the creators of the publicly available Critical Thinking Assessment for Literacy in Visualizations (CALVI) and Massachusetts Visualization (MASSVIS) datasets, which served as the basis for the analyses presented in this study. The author also acknowledges the independent nature of this research, which was conducted without institutional or external support.

Artificial Intelligence (AI) Use Statement: The author confirms that no artificial intelligence (AI) tools were used in the conception, analysis, interpretation, or writing of this manuscript.

Author Contribution: The author was solely responsible for the conception and design of the study, data collection, data analysis, interpretation of the results, and preparation of the manuscript.

Conflict of Interest: The author declares no conflict of interest.

Consent to Publish: The author agrees to the publication of this manuscript in International Journal of Recent Innovations in Academic Research.

Data Availability Statement: The datasets analyzed during this study are publicly available. The Critical Thinking Assessment for Literacy in Visualizations (CALVI) and Massachusetts Visualization (MASSVIS) datasets were obtained from publicly accessible sources cited in the manuscript. No new datasets were generated during this study.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable. This study was based exclusively on publicly available datasets and did not involve direct participation of human subjects.

Informed Consent Statement: Not applicable. This study used publicly available anonymized datasets and did not involve direct interaction with human participants.

Research Content: The research content of the manuscript is original and has not been published elsewhere.

References

1. National Aeronautics and Space Administration. 2011. Human factors in space exploration (NASA Technical Reports Server No. NTRS 20110011268). https://ntrs.nasa.gov/citations/20110011268

2. National Aeronautics and Space Administration. 2025. NASA spaceflight human-system standard: Volume 2: Human factors, habitability, and environmental health (NASA-STD-3001, Rev. E). Section 5.0, Human Performance and Error. https://ntrs.nasa.gov/citations/20250004555
3. McCandless, J.W., McCann, R.S., Berumen, K.W., Gauvain, S.S., Palmer, V.J., Stahl, W.D. and Hamilton, A.S. 2005. Evaluation of the Space Shuttle Cockpit Avionics Upgrade (CAU) displays. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 49(1): 10–14.
4. McCandless, J.W., Hilty, B.R. and McCann, R.S. 2005. New displays for the Space Shuttle cockpit. Ergonomics in Design, 13(4): 15–20.
5. Holden, K., Ezer, N. and Vos, G. 2013. Evidence report: Risk of inadequate human-computer interaction (JSC-CN-28851). National Aeronautics and Space Administration. https://ntrs.nasa.gov/citations/20130014064
6. Padilla, L.M., Creem-Regehr, S.H., Hegarty, M. and Stefanucci, J.K. 2018. Decision making with visualizations: A cognitive framework across disciplines. Cognitive Research: Principles and Implications, 3: 29.
7. Borkin, M.A., Vo, A.A., Bylinskii, Z., Isola, P., Sunkavalli, S., Oliva, A. and Pfister, H. 2013. What makes a visualization memorable? IEEE Transactions on Visualization and Computer Graphics, 19(12): 2306-2315.
8. Borkin, M.A., Bylinskii, Z., Kim, N.W., Bainbridge, C.M., Yeh, C.S., Borkin, D., Pfister, H. and Oliva, A. 2015. Beyond memorability: Visualization recognition and recall. IEEE Transactions on Visualization and Computer Graphics, 22(1): 519-528.
9. Ge, L.W., Cui, Y. and Kay, M. 2023. CALVI: Critical thinking assessment for literacy in visualizations. In: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (Article 815, pp. 1–18). Association for Computing Machinery. https://doi.org/10.1145/3544548.3581406
10. Verma, A., Mukherjee, K., Potts, C., Kreiss, E. and Fan, J.E. 2025. CHART-6: Human-centered evaluation of data visualization understanding in vision-language models. arXiv. https://doi.org/10.48550/arXiv.2505.17202
11. NASA Ames Research Center, Human Systems Integration Division. 2026, March 3. Cockpit display design/Intelligent Spacecraft Interface Systems (ISIS). National Aeronautics and Space Administration. https://www.nasa.gov/human-systems-integration-division/cockpit-display-design-intelligent-spacecraft-interface-systems/
12. Lounis, C., Peysakhovich, V. and Causse, M. 2021. Visual scanning strategies in the cockpit are modulated by pilots' expertise: A flight simulator study. PLoS One, 16(2), e0247061.
13. Peißl, S., Wickens, C.D. and Baruah, R. 2018. Eye-tracking measures in aviation: A selective literature review. The International Journal of Aerospace Psychology, 28(3-4): 98-112.

Citation: Yujeong Yi. 2026. Quantifying the Effect of Misleading Visualizations on Human Judgment with an Explainable Machine-Learning Analysis of Visual Deception and Memorability for Space Interface Design. International Journal of Recent Innovations in Academic Research, 10(3): 82-91.
Copyright: ©2026 Yujeong Yi. This is an open-access article distributed under the terms of the Creative Commons Attribution International License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.