Visualizing Uncertainty with Icon Arrays:
Communicating the Diagnostic Accuracy of a Cancer Screening Test

Tilcia Woodville-Price[1], Peter D. Donnelly[2], Areti Manataki[1]
[1]School of Computer Science, [2]School of Medicine, University of St Andrews, United Kingdom

Motivation

There are a handful of existing studies which look at **incorporating uncertainty into icon arrays** in medical communication.

Conflicting evidence related to comprehensibility and how it impacts decision making.

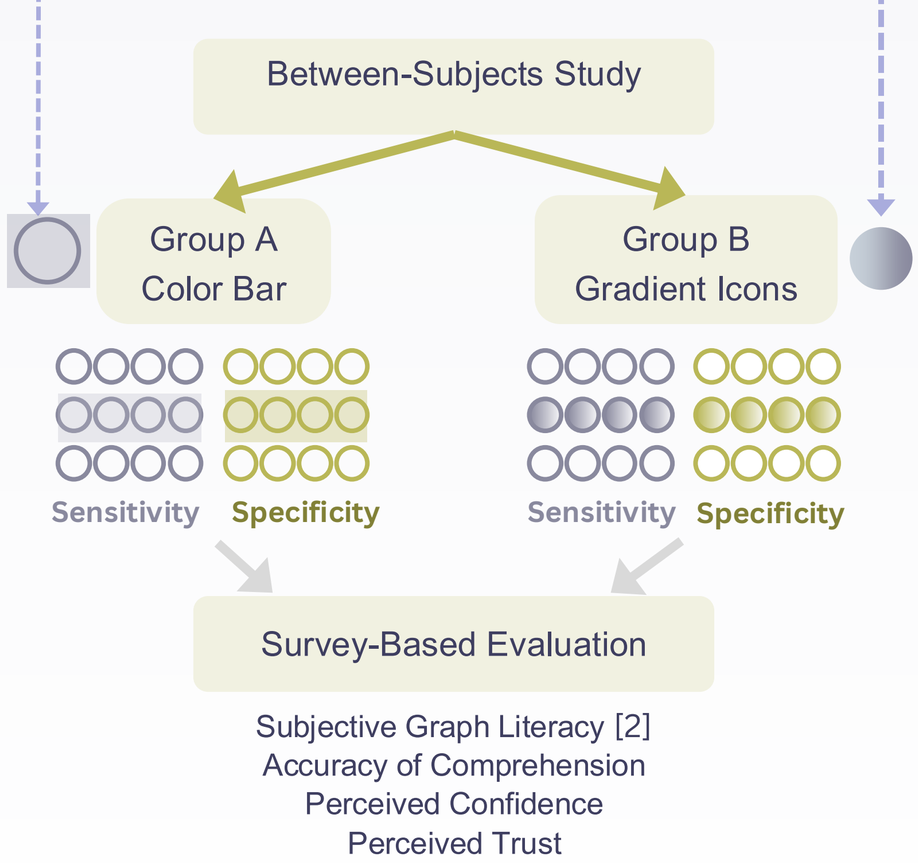
Research Questions

1. Is there a difference in **outcome measures** (accuracy of comprehension, perceived confidence, and perceived trustworthiness) **between visualization techniques** (visualizing uncertainty with color bars and gradients in **icon arrays**)?

2. What is the relationship between **graph literacy** and outcome measures when visualizing uncertainty?

Methods

- **Visualization format**: 10 x 10 static icon arrays
- **Uncertainty**: numerical range, representing diagnostic accuracy (mean and 95% confidence intervals) from a Cochrane Review of collated studies [1].
- **Visual encodings**: 2 visual encodings of uncertainty
  - Color Bar
  - Gradient (within icon)
- **Stimuli**: 2 icon arrays to illustrate the hypothetical screening scenarios:
  - Screened who have cancer (sensitivity*)
  - Screened who do not have cancer (specificity**)

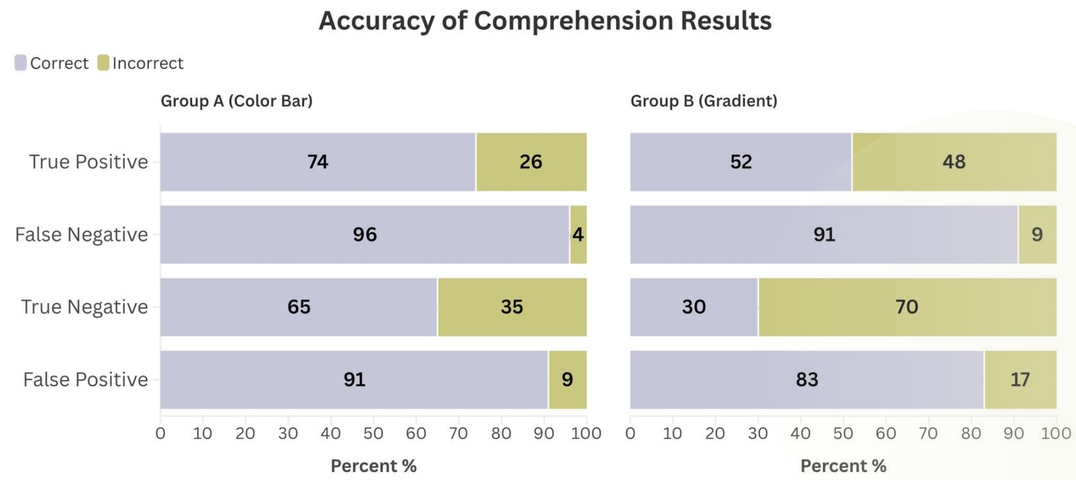


Sensitivity is the ability of a test to yield a positive result for a subject who has that disease. Sensitivity=(True Positives)/(True Positives +False Negatives).
**Specificity* is the ability of the test or instrument to obtain normal range or negative results for a person who does not have a disease. Specificity=(True Negatives)/(True Negatives +False Positives) [3]

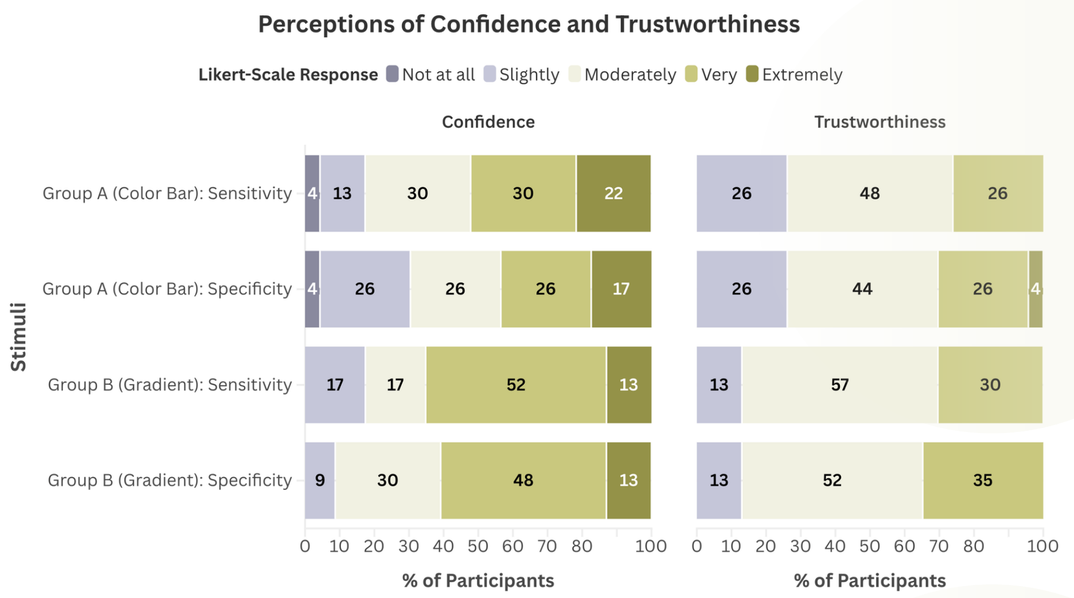
Results

Accuracy of Comprehension*	Subjective Graph Literacy	Perceived Confidence	Perceived Trust
Color Bar > Gradient	Weak correlation with outcome measures	No significant difference between groups	No significant difference between groups
Mean Accuracy: 81.5% vs 64.1%	Mean Score: 49.9 out of 60 across groups	Median composite score: 7 out of 10	Median composite score: 6 out of 10

*Statistically significant result



Color bar encoding resulted in greater accuracy of comprehension than gradient shaded icon (81.5% vs 64.1%).



When presented with uncertainty, participants refrained from extreme perceptions of trustworthiness.

Future Work

- **Qualitative analysis** via semi-structured interviews aiming to look deeper into participant comprehension and perceptions of uncertainty visualization.
- **Data-driven storytelling techniques** [4] may help communicate the multiple outcomes and are also worth exploring.

References

[1] SCHREUDERS E. H., VAN ROON A., VAN DAM L., ZAUBER A. G., LANSDORPVOGELAAR I., BRAMER W., BERHANE S., DEEKS J. J., STEYERBERG E. W., VAN LEERDAM M. E., ET AL.: Guaiac-based fecal occult blood tests versus fecal immunochemical tests for colorectal cancer screening in average-risk individuals. Cochrane Database of Systematic Reviews, 6 (2022). doi:10.1002/14651858.CD009276.pub2.

[2] GARCIA-RETAMERO R., COKELY E. T., GHAZAL S., JOERIS A.: Measuring graph literacy without a test: A brief subjective assessment. Medical Decision Making 36, 7 (2016), 854–867. doi:10.1177/0272989X16655334.

[3] GLAROS A. G., KLINE R. B.: Understanding the accuracy of tests with cutting scores: The sensitivity, specificity, and predictive value model. Journal of clinical psychology 44, 6(1988), 1013–1023. doi:10.1002/1097-4679(198811)44:6<1013::AID-JCLP2270440627>3.0.CO;2-Z.

[4] MEUSCHKE M., GARRISON L. A., SMIT N. N., BACH B., MITTENENTZWEI S., WEISS V., BRUCKNER S., LAWONN K., PREIM B.: Narrative medical visualization to communicate disease data. Computers & Graphics 107 (2022), 144–157. doi:10.1016/j.cag. 2022.07.017.

Acknowledgments

This work is supported by the St Leonard's College World Leading Doctoral Scholarship at the University of St Andrews.
The researchers would like to thank the participants for taking part in this study.

University of St Andrews