International Society of
the Learning Sciences

Supporting Multidimensional Data Analysis for High-School Students in the Era of Machine Learning

Xiaofei Zhou, Jingwan Tang, Qinqin Xiao, Yushan Zhou, Zhen Bai
xzhou50@ur.rochester.edu, jtang21@u.rochester.edu, qxiao5@u.rochester.edu,
yzhou101@ur.rochester.edu, zhen.bai@rochester.edu
University of Rochester

Abstract: Machine Learning (ML) opens exciting scientific opportunities in K-12 STEM classrooms. However, students struggle with interpreting ML patterns due to limited data literacy. Face glyphs offer unique benefit by leveraging our brain's facial feature processing. Yet, they have limitations like lacking contextual information and data biases. To address this, we created three enhanced face glyph visualizations: feature-independent and feature-aligned range views, and the sequential feature inspector. In a study with 25 high school students, feature-aligned range visualization helped contextual analysis, and the sequential feature inspector reduced missing data risks. Face glyphs also benefit the global interpretation of data.

Introduction

Similar to how microscopes reveal the micro world, 21st-century learners can use machine learning (ML) to uncover hidden patterns in multidimensional scientific data. This data-driven discovery aligns with K-12 science practices and crosscutting concepts learning (NGSS, 2013). However, children face challenges in understanding ML patterns due to limited ML knowledge. Exploring multidimensional data globally, rather than focusing on individual attributes, remains a hurdle for K-12 learners (Ben-Zvi & Arcavi, 2001). Face glyphs visually map individual attributes to facial features, like eye radius (Chernoff, 1973). They offer unique benefits for children analyzing ML-revealed patterns due to how our brains process facial features and glyphs' affordance for superposition comparison & direct manipulation. However, face glyphs have two main limitations. First, they often lack an obvious connection between visual features and data context (attribute and value range). To address this, we propose feature-aligned face scales and feature-independent bar scales. Second, distinct facial features may lead to attention biases toward different data attributes. To mitigate this, we designed a dial gadget based on guided search theory (Healey & Enns, 2012), forcing students to explore all facial features sequentially. We evaluated these enhanced face glyphs against a segment-style star glyph with 25 high school students.

Theoretical framework

Existing research demonstrated that children have the potential to apply basic ML methods to discover scientific knowledge from ML-revealed multidimensional patterns (Wan et al., 2020). We chose k-means clustering to mine the patterns for students to analyze for two reasons. First, cluster analysis is widely used in the science community to reveal knowledge (Romesburg, 2004), and thus offer diverse curriculum-aligned contexts. Second, k-means clustering, with its focus on multidimensional data, similarity computation, and centroids, serves as a foundational ML concept. This knowledge can transfer to other ML methods.

Challenges in supporting global interpretation of patterns revealed by ML. This involves identifying, describing, and explaining multidimensional patterns and trends (Ben-Zvi & Arcavi, 2001). It is challenging for children to understand multiple factors that jointly affect an outcome (Kuhn et al., 2015), and to read data graphs and tables (Shah & Hoeffner, 2002). Glyphs encode multiple attributes with visual features accessible to novice learners (Ward, 2015) and utilize perceptual organization principles to assist efficient information processing.

Unique benefits afforded by face glyph to tackle the challenges. First, visualization of faces is particularly attractive to children (Ryokai et al., 2012). Being fun and unusual, face glyphs may stimulate low-performance students' self-efficacy and intrinsic enjoyment (Moyer, 2001). Second, with the human brain's special processing of facial features, children may process global patterns in an accessible way, potentially better than adults (Tsurusawa et al., 2008). Third, face glyphs outperformed other iconic glyphs (e.g., stars) in similarity comparison (Wilkinson, 1982), which is fundamental in ML-supported pattern mining. However, face glyphs have two major limitations. First, the context of data attributes and values, which is critical for data-driven scientific inquiry (Valero-Mora & Ledesma, 2011), is absent from existing iconic glyphs. Second, it is critical to have a global view of multidimensional data without missing any interesting attributes, especially for exploratory analysis. Glyphs might cause biases in recognizing attributes--people pay more attention to eyes than noses (Ward et al., 2015).

Integrate context information. Little effort has been spent on integrating the dataset value context (i.e., the data range) into face glyphs. The superposition comparative visualization approach is efficient in minimizing attention switches in data analysis during exploratory visual analysis. This suggests a feature-aligned visualization--displaying data range by the same graphical mark (Thompson et al., 2021). When learners interact with the same graphical marks, there is no attention switch from facial features to another visual form. This leads to one less step than using a different graphical feature.

Support less biased feature inspection. To support a less biased visual search of identifying features of interest in glyphs, we utilize the guided search theory, which decomposes visual search into bottom-up and top-down activation processes (Healey & Enns, 2012). Bottom-up activation measures the differences between a visual element and its neighbors, while top-down activation is driven by users intentionally searching for the visual information needed. The theory hypothesized that the attention will be drawn to peaks in the image area with the largest combination of bottom-up and top-down activation. This suggests an interactive element guiding learners to sequentially distribute their attention across different attributes more equally. Sequential guidance of learners' attention may also reduce cognitive load during analysis. Fixed-order inspection reduced the working memory needed for what features they have already considered previously (Shneiderman et al., 2016).

Figure 1

Face glyph design elements and the segment-style star glyph.

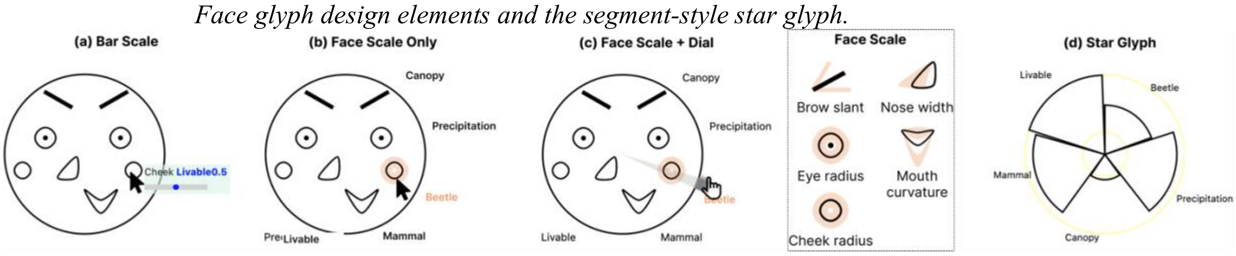
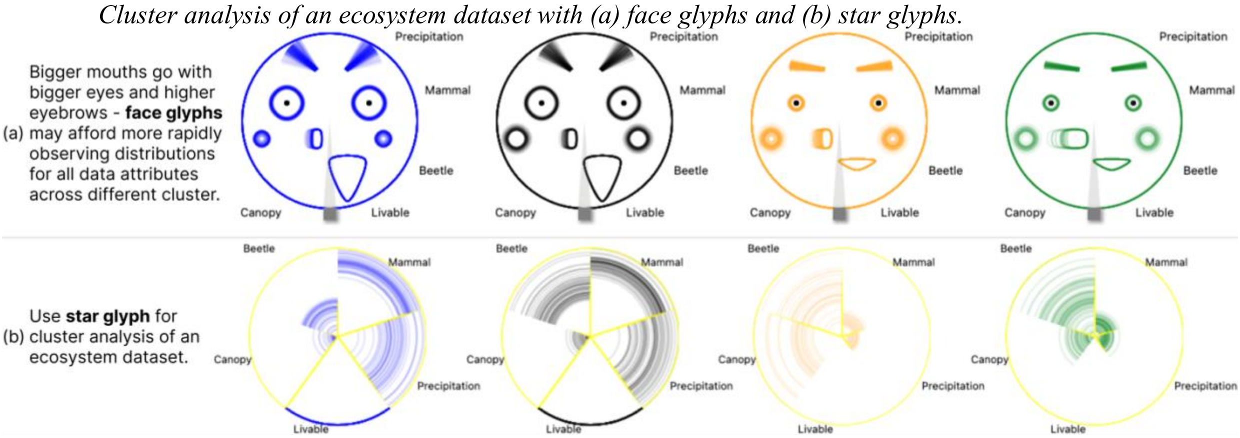


Figure 2

Cluster analysis of an ecosystem dataset with (a) face glyphs and (b) star glyphs.



Glyph design for global interpretation of data

Face Glyph Enhancements

To encode data attributes, we selected facial features including eyebrow slant degree, eye radius, cheek size, mouth curvature, and nose width (Fig. 1). In the overlay view for a group of data, face glyphs turn semi-transparent and are stacked together. Learners utilize color saturation, and contour thickness (i.e., facial feature variation) to analyze the value range and distribution of a certain group of data (Fig. 2).

Two scale designs for range visualization. First, face scales provide feature-aligned range visualization (Fig. 1 Face scale). Once it is activated by cursor hovering (Fig. 1.b) or by a dial gadget (Fig. 1.c), the data attribute and value range are highlighted. Second, displaying bar scales off glyphs supports the overview + detail scheme (Cockburn et al., 2009). Learners first form an overview of patterns and then look into the details. As feature-independent, bar scales will pop up when learners hover their cursor over individual facial features (Fig. 1.a).

A dial gadget as the sequential feature inspector. This forces learners to browse all the data attributes in sequential order. This element looks like a clock hand and triggers the context information (data attribute and range visualization). When the dial is dragged over a certain facial feature (e.g., cheek size), the context information will be triggered on the glyph. When a learner needs more detailed information to analyze a certain data attribute, they need to drag the dial to scan through all the rest of the data attributes. This could mitigate people's potential attention differences across different facial features. The dial aims to guide learners away from fixating on a single glyph feature and to reduce biases in analysis. By providing a mechanism for sequentially exploring different attributes, the dial design encourages learners to take a more holistic view of the data.

Star glyph design for comparison

Star glyphs, commonly compared with face glyphs in empirical studies, mainly encode the value either as a line length or as a segment length. Star glyphs carry no extra distracting semantic meanings and are closer to conventional charts in K-12, such as bar charts and pie charts. We chose segment style (Fig. 1.d) as it also supports superposition comparative visualization (Fig. 2.b). To make an attribute with minimal value visible in the overlay view, the start of each data attribute is offset from the center to avoid ambiguity. For range visualization, we added two yellow circles to represent the min and max for each attribute, using the same graphical mark as the glyph visual feature. For less biased feature inspection, with identical visual channels encoding data, star-based glyphs afford uniformity across different attributes.

Research method

Research participants and procedure. Following our approved Institutional Review Board protocol, we recruited 25 participants (10 females, 14 males, and one preferred not to disclose their gender). Their grade levels ranged from seven to 11 (Mean = 9.76, SD = 1.16). 16 students had no prior AI experience, five were familiar with AI-related products, and four had basic coding knowledge. The study was conducted online via Zoom, with each session lasting about 1.5 hours. Participants experienced all design conditions in different sequences to minimize learning effects. We used partial counterbalancing and the Latin square method to determine condition combinations for four participant groups. Four datasets were rotated across the design options. A pre-survey collected students' background information. Then participants were asked to "think aloud" while going through a set of data analysis tasks independently. The researcher reminded participants to keep verbalizing their actions and reasoning process when they fell into silence.

Data analysis. We tested whether data interpretation performance differed by glyph designs through within-subject analysis. The data was not normally distributed based on the Shapiro-Wilk test. We used a non-parametric test, the Friedman test, on each measure to assess differences between the four conditions. For further pairwise comparison, we applied the Wilcoxon signed-rank test to each combination. The Bonferroni adjusted-p = .01(0.05/6) is used to avoid the Type I error in multiple comparisons.

Results

Support global interpretation of multidimensional data for children

Face glyphs outperformed star glyphs in the accuracy of global interpretation. The Friedman test revealed significant differences across different glyph designs ($X^2(3) = 16.50$, $p < .001$). Among all four glyphs, star glyphs performed the worst. Post hoc analysis using the Wilcoxon test showed that face glyphs with face scales only ($p = .001$) and dial & face scales ($p < .001$) significantly outperformed star glyphs. When analyzing patterns, face glyph users noticed similarities within clusters and differences between groups. In contrast, star glyph users tended to focus on local features, ignoring between-group differences. Face glyph users identified facial features that split clusters, interpreted intra-group patterns, and compared differences between cluster groups.

Support analyzing data in the context

We compared face glyphs with face scales and bar scales. The Friedman test revealed significant differences in students' performance for contextualizing analysis with both data attributes ($X^2(3) = 27.69$, $p < .001$) and data range ($X^2(3) = 43.00$, $p = .007$). Post hoc analysis showed that face scales significantly outperformed bar scales ($p < .001$) for data attribute context. Additionally, face glyphs with face scales performed better than face glyphs with bar scales ($p = .006$ and .007) for data range context.

Sequential feature inspector for less biased feature inspection

We compared face glyphs with and without the dial for sequential feature inspection. The Friedman test revealed significant differences across different glyph designs ($X^2(3) = 10.32$, $p = .02$). Post hoc analysis using the Wilcoxon test showed that face glyphs with a dial performed significantly better in inspecting individual data attributes equally compared to face glyphs with bar scales ($p = .008$), face scales only ($p = .009$), and star glyphs ($p = .04$). Interestingly, not all learners utilized the dial inspection function consistently: eight used it for all tasks, 12 for half or less, and three not at all. Further analysis showed significantly less bias ($t(153) = 2.46$, $p = .01$) in tasks where participants used the dial.

Conclusion and Discussion

To support children in analyzing patterns revealed by k-means clustering, we proposed four glyph designs. A study with 25 high school students showed that the sequential feature inspector (the dial) effectively supported less biased interpretation of data attributes, and the feature-aligned range visualization (face scales) is effective for contextualized analysis. Many iconic glyphs lack data range information. Feature-aligned range visualization can benefit other iconic glyphs (e.g., cars, flowers). The dial successfully mitigated face glyphs' limitations of biased feature inspection. This confirms the effectiveness of utilizing guide search theory for more effective data analysis and offloading short-term memory by forcing sequential inspection. Not all participants fully utilized the dial, and this suggests adapting the design based on different learner styles and making the dial interaction more inviting and rewarding. Furthermore, face glyphs supported the global interpretation of multidimensional patterns better than star glyphs. Global interpretation of data involves recognizing relationships across multiple attributes, requiring memorization of data details. Face glyphs outperform star glyphs, possibly because facial features aid recognition and memorization (Applegate & Cohen, 2017). Face glyphs offer variability in expressing data, enhancing pattern visibility (Fazlagić et al., 2018).

Acknowledgment

We would like to acknowledge the support from the National Science Foundation RETTL program (Award No. 2225227) and the industry fellowship through the NSF NRT project (Award No. 1922591).

References

Applegate, M., & Cohen, J. (2017). Communicating graphically. Cultural Politics an International Journal, 13(1).

Ben-Zvi, D., & Arcavi, A. (2001). Junior high school students' construction of global views of data and data representations. Educational Studies in Mathematics, 45(1), 35–65.

Chernoff, H. (1973). The Use of Faces to Represent Points in k-Dimensional Space Graphically. Journal of the American Statistical Association, 68(342), 361–368.

Cockburn, A., Karlson, A., & Bederson, B. B. (2009). A review of overview+detail, zooming, and focus+context interfaces. ACM Computing Surveys, 41(1), 1–31.

Fazlagić, J., Loopesko, W., Matuszak, L., & Johnson, R. (2018). Perspectives and Good Practices in Visualization of Knowledge About Public Entities. In Information Visualization Techniques in the Social Sciences and Humanities (pp. 195–213). IGI Global.

Healey, C. G., & Enns, J. T. (2012). Attention and visual memory in visualization and computer graphics. IEEE Transactions on Visualization and Computer Graphics, 18(7), 1170–1188.

Kuhn, D., Ramsey, S., & Arvidsson, T. S. (2015). Developing multivariable thinkers. Cognitive Development.

Moyer, P. S. (2001). Are We Having Fun Yet? How Teachers Use Manipulatives to Teach Mathematics. Educational Studies in Mathematics, 47(2), 175–197.

National Research Council. 2013. Next generation science standards: For states, by states. (2013).

Romesburg, C. (2004). Cluster Analysis for Researchers. Lulu.com.

Ryokai, K., Raffle, H., & Kowalski, R. (2012). StoryFaces: pretend-play with ebooks to support social-emotional storytelling. Proceedings of the 11th International Conference on Interaction Design and Children.

Shah, P., & Hoeffner, J. (2002). Review of Graph Comprehension Research: Implications for Instruction. Educational Psychology Review, 14(1), 47–69.

Shneiderman, B., Plaisant, C., Cohen, M., Jacobs, S., Elmqvist, N., & Diakopoulos, N. (2016). Designing the User Interface: Strategies for Effective Human-Computer Interaction (6th ed.). Pearson.

Thompson, J. R., Liu, Z., & Stasko, J. (2021). Data Animator: Authoring Expressive Animated Data Graphics. Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, Article Article 15.

Tsurusawa, R., Goto, Y., Mitsudome, A., Nakashima, T., & Tobimatsu, S. (2008). Different perceptual sensitivities for Chernoff's face between children and adults. Neuroscience Research, 60(2), 176–183.

Valero-Mora, P. M., & Ledesma, R. D. (2011). Using interactive graphics to teach multivariate data analysis to psychology students. Journal of Statistics Education, 19(1).

Wan, X., Zhou, X., Ye, Z., Mortensen, C. K., & Bai, Z. (2020, June). SmileyCluster: supporting accessible machine learning in K-12 scientific discovery. In Proceedings of the interaction design and children conference (pp. 23-35).

Ward, M. O., Grinstein, G., & Keim, D. (2015). Interactive Data Visualization: Foundations, Techniques, and Applications, Second Edition. CRC Press.

Wilkinson, L. (1982). An experimental evaluation of multivariate graphical point representations. Proceedings of the 1982 Conference on Human Factors in Computing Systems, 202–209.