# UC Santa Cruz
## UC Santa Cruz Previously Published Works

**Title**

Large-Scale Analysis of Visualization Options in a Citizen Science Game.

**Permalink**

https://escholarship.org/uc/item/7xt5r0jh

**Authors**

Miller, Josh
Lee, Vivian
Cooper, Seth
et al.

**Publication Date**

2019-10-01

**DOI**

10.1145/3341215.3356274

Peer reviewed

# Large-Scale Analysis of Visualization Options in a Citizen Science Game

**Josh Aaron Miller**,
Northeastern University Boston, MA, USA

**Vivian Lee**,
Northeastern University Boston, MA, USA

**Seth Cooper**,
Northeastern University Boston, MA, USA

**Magy Seif El-Nasr**
Northeastern University Boston, MA, USA

## Abstract

Visualization is a valuable tool in problem solving, especially for citizen science games. In this study, we analyze data from 36,351 unique players of the citizen science game Foldit over a period of 5 years to understand how their choice of visualization options are affected by expertise and problem type. We identified clusters of visualization options, and found differences in how experts and novices view puzzles and that experts differentially change their views based on puzzle type. These results can inform new design approaches to help both novice and expert players visualize novel problems, develop expertise, and problem solve.

## Keywords

citizen science game; visualization; expertise

## Introduction

Visualization is a valuable tool in problem solving [6]. Although previous research has demonstrated a connection between expertise and visualization approaches [18], we still don't understand the generalized process of how visualization aids problem solving and how expertise affects one's visualization techniques.

Experts and novices vary in many ways with respect to problem solving, including what information they store in memory, their attitudes toward the problem, their strategies applied, and their visualization techniques [1]. For example, one study found that expert

mathematicians rely on visualizations to guide their planning and execution more than novices [19].

The process of visualizing a problem in order to solve it is especially practical for citizen science games, in which citizen scientists attempt to solve creative, difficult, complex, and novel problems through the power of crowdsourcing [4]. Visualization in citizen science games is also important because players cite the visual appeal as a key factor in their motivation for playing [5, 20].

Improper visualizations can make problem solving more difficult by adding perceptual or cognitive load. Perceptual load can be thought of as clutter: irrelevant information on-screen can distract the user and reduce their attention and awareness toward task-relevant processing [8]. Additionally, more complex information representations can increase cognitive load and subsequently decrease task performance if it exceeds the user's working memory limitations [2].

Additionally, previous research has explored how a user interface conveys information in games, such as through sound [14] and "immersive" (i.e., diegetic) displays [9]. Researchers have also recently begun exploring the user experience of data visualizations in general [16].

Finally, the value of understanding how problems are visualized in citizen science games is that the games can be customized and adjusted to the needs of the player. Most commercial games offer customizable visual options of some kind, such as adding subtitles or toggling the visibility of certain user interface elements. In the citizen science game EyeWire[1], players can even customize their roles: after earning a particular rank, they may opt into new abilities which let them fundamentally alter how they play the game. From colorblind modes to advanced user interfaces, games let their players change how they view and approach their tasks. **Understanding how and why players choose to visualize their games can inform better user-centered visual design.**

In this paper, we examine Foldit [3], a citizen science puzzle game about protein folding, to explore how players visualize novel problems. The importance of visualization for this class of problem is well-known, as illustrated by the existence of graphical tools for molecular visualizations (similar to that of Foldit), such as PyMOL [17] and UnityMol [10]. We consider the following research questions:

RQ1. How do players view Foldit puzzles?

RQ2. Do experts and novices view puzzles differently?

RQ3. Do players view puzzles differently based on the type of problem?

---

[1]https://eyewire.org/

## Methodology

### Data Collection

Gameplay data were collected[2] from 72,041 unique players of Foldit[3] between June 2011 and August 2016. Each log was collected automatically at the end of each game session when the player exited the game. In total, 578,191 logs were collected across 1,973 unique puzzles. We cleaned our dataset by removing corrupted logs and logs that came from tutorial or practice ("Beginner") puzzles. After this filtering process, there remained 439,469 logs from 36,351 unique players with a mean of 12.09 (SD = 139.44) logs per player.

Each data log contained meta-data (e.g. anonymized player ID and timestamp) as well as data on the visualization option settings themselves (henceforth called the player's "view"). There were 46,509 unique views. A summary of the relevant options used by players are described in Table 1. We operationally defined expert players as those who earned a high score in at least two unique puzzles, where a high score is defined as being in the top 5% of all scores for that puzzle. Although choosing two high scores is an arbitrary distinction, this resulted in defining only 1,007 experts (2.8% of the original user population), yet the experts' collective data make up 71.0% of all post-processed logs. This is intuitive, since most experts play very frequently, while most novices may only interact with the game a few times. For all analyses comparing experts to novices (i.e., non-experts), we randomly selected an equal number of novices (n=1,007) from novices who had at least one log of each puzzle category analyzed (Design, Prediction, Electron Density, and Hand-Folding). These same novices were used in all expert/novice comparisons.

### Clustering

To answer RQ1[4], we applied agglomerative clustering[5] to the 46,509 unique views using Euclidean distance and Ward linkage. Because this technique gives weight to the number of distinct views rather than the popularity of views, we see artifacts such as all clusters having "advanced mode" enabled, since this option allows more distinct views to be set than the few options available when advanced mode is disabled. However, this approach allowed us to compare view settings more fairly. We considered only unique views to avoid biasing toward either expert or novice data and instead clustered by the similarity of the views themselves. We found that experts contributed many more unique views, yet this is a finding rather than a bias.

In order to determine how many clusters to use, we visualized the dendrogram of the data and plotted the average sum of squared errors per cluster by the number of clusters. Through these methods, we determined 6 clusters to be the most reasonable grouping; these clusters provide an answer to RQ1, expressing the most central ways of viewing the puzzles.

As shown in Figure 1, Cluster 1 captures all default settings, so we interpret this as the Default cluster. Cluster 2, Minimalist, seems to capture a pared-down view, which may

---

allow the player to focus on the secondary structure of the protein without the added complexity of the sidechains. Cluster 3, Show Sidechains, is the reverse of minimalism, showing much more detail in how the sidechains are organized. Although it was not used very often, this cluster seems to represent the middle ground between Clusters 1 and 5. Cluster 4, Color Coded, is unique in using both CPK coloring and relative score coloring: none of the other clusters use either of these color schemes. Cluster 5, Show Everything, is the most detailed, and seems to be useful for identifying issues such as exposed hydrophobics[6]. Finally, Cluster 6, Dark Minimalist, seems to mirror Cluster 2 but with a thinner structure and darker background.

## Results

To answer RQ2[7], we compared the frequencies of views in each cluster[8] between expert and novice (i.e., non-expert) players. For each player, we counted how many of their views were in each cluster, then normalized this count to a weight of exactly 1. For example, if a player had 2 views in Cluster 1 and 3 views in Cluster 4, we counted this as frequency weights of 0.4 and 0.6, respectively. In this way, all players contributed an equal proportion to how the frequencies were calculated. Finally, we performed a chi-square test of goodness of fit to determine whether experts and novices used similar views. The groups were significantly different, $\chi^2$ (5, N = 2,014) = 1,578.80, $p < 0.001$. As shown in Table 2, this result is driven by novices primarily using clusters 1 and 5, whereas experts tend to diversify across clusters.

To answer RQ3[9] we further analyzed the difference between experts and novices by four puzzle categories: Design, Prediction, Electron Density, and Hand-Folding. To determine whether puzzle category had an effect on the player's views, we performed multiple chi-square tests of independence on the frequency of views in a contingency table. For experts and novices, we first ran an omnibus test across all puzzle categories; if the test was significant, we ran post-hoc tests comparing each puzzle type to all others. We applied Bonferroni corrections in these post-hocs to account for multiple comparisons. For novices, no significant effects of puzzle category were detected (see Table 3). For experts, the omnibus test was significant ($\chi^2$ (5, N = 1,007) = 49.94, $p < 0.001$). Post-hoc analyses showed that both Prediction ($\chi^2$ (5, N = 1,007) = 15.28, $p < 0.01$) and Hand-Folding ($\chi^2$ (5, N = 1,007) = 24.34, $p < 0.001$) puzzles were significantly different than expected after correction (see Table 4). Differences in Prediction seem to be driven by an increase in the frequency of views being in Clusters 1 and 5 and a decrease in the frequency of remaining clusters, while differences in Hand-Folding were driven by the opposite: primarily an increase in the frequencies of Clusters 2, 3, and 4 and a decrease in the rest.

---

[6]These are represented as yellow bubbles on the protein structure.
[7]Do experts and novices view puzzles differently?
[8]For this analysis, we consider view settings by cluster instead of analyzing raw option variables for two reasons. First, we can more easily analyze differences in post-hoc analyses. Second, several options are rarely or never used, complicating statistical analysis.
[9]Do players view puzzles differently based on the type of problem?

## Discussion

By applying hierarchical agglomerative clustering to all of the unique views in our dataset, we identified 6 clusters to represent the different views of the players analyzed (Figure 1). Cluster 1 (Default Settings) captured over half of all novice data since novices are less likely to change their view settings. Clusters 2 (Minimalist) and 6 (Dark Minimalist) seem especially useful for earning a high score: high scoring views use Cluster 2 between 24–34% of the time and Cluster 6 between 18–22% of the time for the four main puzzle categories analyzed.

Additionally, we contacted five expert players about their view preferences. Overall, their comments were of the form "when attempting task X, I use option Y" which supports our findings that experts (1) use a wider variety of views and (2) differentially adjust their views to the task. Indeed, they explicitly mentioned that view settings are context-dependent. Notably, four experts expressed that they have a "default" way of viewing the puzzle when they aren't using a task-specific view, three of which were described in form and function as matching the Minimalist clusters. The fourth expert's default setting resembled the "Show Everything" cluster, albeit with the addition of CPK coloring, which the expert removes for "rough drafts" of a puzzle solution. Other task-specific preferences were expressed, such as aligning sheets (Show bondable atoms) and managing disulfide bonds (there were a variety of responses for how they approach this task). Additionally, one expert requested new visualization features and another requested more hotkey shortcuts to access toggling these options, further suggesting that experts make heavy use of these features and that customizable view options support task performance. Interestingly, none of the experts expressed a need for "preset" settings to toggle between popular use cases. This is perhaps because their needs are more nuanced than presets would allow.

To these experts, there are many types of problems, both general (shape of the protein, the formation of hydrogen bonds) and specific (viewing ligands or electron density clouds). Details on how expert players assess these puzzles in Foldit has also been researched by Ponti et al. [13].

Future work can examine how visualization preferences are impacted by the existence of communities of practice [21] or knowledge communities [11]. Specifically, when individuals attempt problem solving as a group, collaboration may either lead to sharing resources and converging on knowledge or divergence and specialization among collaborators [7]. The extent to which the group collaborates is based on the created "joint problem space" [15]. Therefore, it would be worthwhile to investigate how players specialize or collaborate in the way they view the problem.

### Design Considerations

From these results, we can conclude that experts solving a complex visual task use and require customizable visualization options. However, for novices these customizations may be cognitively out of reach. Although none of the expert players expressed a need for preset views, such groupings (e.g., our clusters) may be helpful for introducing novices to a wider array of views as part of their paths to expertise. Moreover, to the untrained eye, the value of

each visualization option is unclear. One solution would be to add tooltips on hover to describe possible use cases of each option. Finally, the game could define more helpful default settings based on the user's activity.

For experts, their primary needs are (1) additional autonomy and flexibility and (2) reduced effort for frequent inputs. In commercial games, this need is most often resolved by an active modding community [22].

### Limitations

The major limitation of this study was that it analyzed only one game's data. Future work may extend this analysis beyond Foldit to other games with customizable visualization options to more fully understand the effects of customizing visualizations. We also acknowledge that in this analysis we are assuming that the players make use of the visualization settings they have enabled (or disabled). For example, when the view is composed of default settings, the player may be simply working with the visualization being offered rather than choosing that view to make use of the information it provides. Additionally, some tutorial levels override these default settings to demonstrate features of the game, which may explain why many novices were using views in Cluster 5, even if they did not change those settings themselves. Although we excluded data from tutorial levels, they may have impacted the default settings of some players which were not explicitly accounted for in our analysis.

## Conclusion

In this study, we analyzed the view options settings of Foldit players to understand how players' choice of visualizations are affected by expertise and problem type. We found significant differences in how experts view puzzles compared to novices and that experts differentially change their view based on puzzle type. Through input from expert players, we qualitatively support these findings from the human perspective. We discuss how these results can inform design changes for the needs of both novices (e.g. guided customization) and experts (e.g. mod support).

## Acknowledgements

## REFERENCES

1. Adams Jonathan Paul, Kaczmarczyk Stefan, Picton Phil, and Demian Peter. 2009. Problem solving and creativity in engineering: Perceptions of novices and professionals. In World Congress on Engineering and Computer Science, San Francisco, USA.

2. Anderson Erik W, Potter Kristin C, Matzen Laura E, Shepherd Jason F, Preston Gilbert A, and Silva Cláudio T. 2011. A user study of visualization effectiveness using EEG and cognitive load. In Computer graphics forum, Vol. 30. Wiley Online Library, 791–800.

3. Cooper Seth, Khatib Firas, Treuille Adrien, Barbero Janos, Lee Jeehyung, Beenen Michael, Leaver-Fay Andrew, Baker David, Popovi Zoran, and players Foldit. 2010. Predicting protein structures with a multiplayer online game. Nature 466, 7307 (2010), 756. [PubMed: 20686574]

4. Curtis Vickie. 2014. Online citizen science games: Opportunities for the biological sciences. Applied & translational genomics 3, 4 (2014), 90–94. [PubMed: 27294021]

5. Curtis Vickie. 2015. Motivation to participate in an online citizen science game: A study of Foldit. Science Communication 37, 6 (2015), 723–746.

6. Downing Ricard E, Moore Joi L, and Brown Steven W. 2005. The effects and interaction of spatial visualization and domain expertise on information seeking. Computers in Human Behavior 21, 2 (2005), 195–209.

7. Fischer Frank, Bruhn Johannes, Gräsel Cornelia, and Mandl Heinz. 2002. Fostering collaborative knowledge construction with visualization tools. Learning and Instruction 12, 2 (2002), 213–232.

8. Lavie Nilli. 2005. Distracted and confused?: Selective attention under load. Trends in cognitive sciences 9, 2 (2005), 75–82. [PubMed: 15668100]

9. Llanos Stein C and Jørgensen Kristine. 2011. Do players prefer integrated user interfaces? A qualitative study of game UI design issues. In Proceedings of DiGRA 2011 Conference: Think Design Play.

10. Lv Zhihan, Tek Alex, Da Silva Franck, Empereur-Mot Charly, Chavent Matthieu, and Baaden Marc. 2013. Game on, science-how video game technology may help biologists tackle visualization challenges. PloS One 8, 3 (2013), e57990.

11. Paavola Sami, Lipponen Lasse, and Hakkarainen Kai. 2004. Models of innovative knowledge communities and three metaphors of learning. Review of educational research 74, 4 (2004), 557–576.

12. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, and Duchesnay E. 2011. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research 12 (2011), 2825–2830.

13. Ponti Marisa, Stankovic Igor, Barendregt Wolmet, Kestemont Bruno, and Bain Lyn. 2017. Chefs Know More Than Just Recipes: Professional Vision in a Citizen Science Game. SocArXiv. 3 20 (2017).

14. Robb James, Garner Tom, Collins Karen, and Nacke Lennart E. 2017. The impact of health-related user interface sounds on player experience. Simulation & Gaming 48, 3 (2017), 402–427.

15. Roschelle Jeremy and Teasley Stephanie D. 1995. The construction of shared knowledge in collaborative problem solving. In Computer supported collaborative learning. Springer, 69–97.

16. Saket Bahador, Endert Alex, and Stasko John. 2016. Beyond usability and performance: A review of user experience-focused evaluations in visualization. In Proceedings of the Sixth Workshop on Beyond Time and Errors on Novel Evaluation Methods for Visualization. ACM, 133–142.

17. Schrödinger LLC. 2015. The PyMOL Molecular Graphics System, Version 1.8. (11 2015).

18. Stylianou Despina A. 2002. On the interaction of visualization and analysis: the negotiation of a visual representation in expert problem solving. The Journal of Mathematical Behavior 21, 3 (2002), 303–317.

19. Stylianou Despina A and Silver Edward A. 2004. The role of visual representations in advanced mathematical problem solving: An examination of expert-novice similarities and differences. Mathematical thinking and learning 6, 4 (2004), 353–387.

20. Tinati Ramine, Luczak-Roesch Markus, Simperl Elena, and Hall Wendy. 2016. Because science is awesome: studying participation in a citizen science game. In Proceedings of the 8th ACM Conference on Web Science. ACM, 45–54.

21. Wenger Etienne C and Snyder William M. 2000. Communities of practice: The organizational frontier. Harvard business review 78, 1 (2000), 139–146. [PubMed: 11184968]

22. Wharton Alexander and Collins Karen. 2011. Subjective measures of the influence of music customization on the video game play experience: A pilot study. Game Studies: the international journal of computer game research 11, 2 (2011).

**CCS Concepts**

- •**Human-centered computing** → **Visualization;** *Graphical user interfaces; User interface design;*
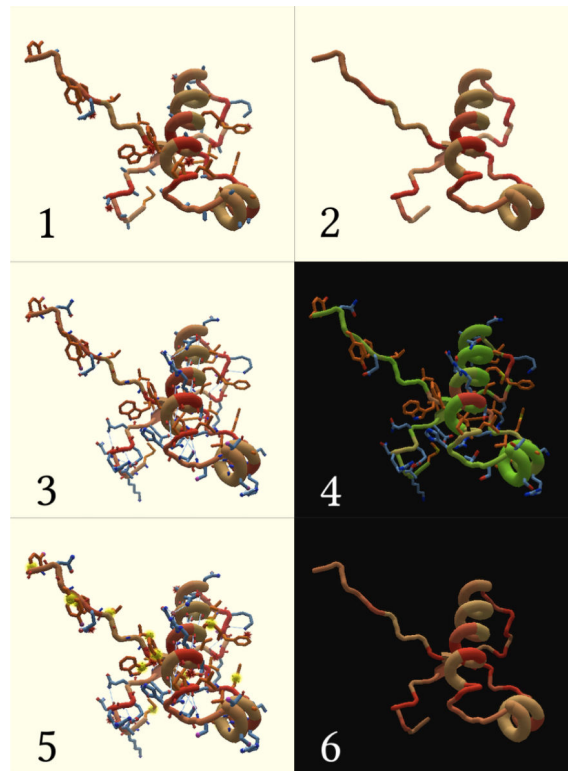
**Figure 1:**
Visualization of View Option Clusters. Each screenshot represents the centroid of a cluster from the agglomerative clustering. To visualize the centroid, we selected the nearest option value (i.e., by rounding binary options and selecting the majority value for categorical options). We interpret these clusters as: (1) Default Settings; (2) Minimalist; (3) Show Sidechains; (4) Color Coded; (5) Show Everything; (6) Dark Minimalist.

**Table 1:**

Summary of View Options used by Players. Options not related to views, not changed from the default, or not visibly different in our clustering analysis are not shown here. We included 61 view options total in our analysis. Default options are marked with an asterisk (*) when first shown on the table. We interpret these clusters as: (1) Default Settings; (2) Minimalist; (3) Show Sidechains; (4) Color Coded; (5) Show Everything; (6) Dark Minimalist.

| Option | C1 | C2 | C3 | C4 | C5 | C6 |
|---|---|---|---|---|---|---|
| Show sidechains | Stubs only* | No | Yes | Yes | Yes | No |
| Dark background | No* | No | No | Yes | No | Yes |
| CPK coloring | No* | No | No | Yes | No | No |
| Relative score coloring | No* | No | No | Yes | No | No |
| Thin rendering style | No* | No | Yes | Yes | No | Yes |
| Show bondable atoms | No* | No | Yes | No | Yes | No |
| Show clashes | Yes* | No | No | No | Yes | No |
| Show issues | No* | No | No | No | Yes | No |
| Show hydrogen bonds | Yes* | No | Yes | No | Yes | No |
| Pulse when working | Yes* | No | Yes | No | Yes | No |

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 2:**

Percentage of View Frequencies by Cluster. Percents are rounded to the nearest integer for readability.

|        | Experts | Novices |
|--------|---------|---------|
| **C1** | 41      | 54      |
| **C2** | 17      | 5       |
| **C3** | 7       | 3       |
| **C4** | 5       | 1       |
| **C5** | 22      | 35      |
| **C6** | 9       | 1       |

**Table 3:**

Percentage of Novice Views by Cluster and Puzzle Category. Percents are rounded to the nearest integer for readability.

|  | Design | prediction | Electro Density | Hand -Folding |
|---|---|---|---|---|
| **C1** | 55 | 50 | 51 | 48 |
| **C2** | 5 | 6 | 8 | 8 |
| **C3** | 3 | 3 | 3 | 4 |
| **C4** | 1 | 1 | 2 | 2 |
| **C5** | 35 | 38 | 35 | 38 |
| **C6** | 1 | 2 | 1 | 1 |

**Table 4:**

Percentage of Expert Views by Cluster and Puzzle Category. Percents are rounded to the nearest integer for readability.

|     | Design | prediction | Electro De | Hand -Folding |
|-----|--------|------------|------------|---------------|
| C1  | 34     | 38         | 33         | 30            |
| C2  | 19     | 18         | 22         | 23            |
| C3  | 9      | 8          | 7          | 11            |
| C4  | 7      | 5          | 7          | 11            |
| C5  | 20     | 22         | 17         | 15            |
| C6  | 11     | 10         | 13         | 10            |