Scaffolding Design to Bridge the Gaps between Machine Learning and Scientific Discovery for K-12 STEM Education

Xiaofei Zhou
University of Rochester
Rochester, United States
xzhou50@ur.rochester.edu

Kaixin Li
University of Rochester
Rochester, United States
kli28@u.rochester.edu

Abdul Moid Munawar
University of Rochester
Rochester, United States
amunawar@u.rochester.edu

Zhen Bai
University of Rochester
Rochester, United States
zhen.bai@rochester.edu

ABSTRACT

Machine Learning (ML) can provide an advanced lens for K-12 students to get their hands on intriguing patterns from real-world data and has the potential to empower young learners with more challenging cognitive skills needed for iterative scientific investigation. However, few efforts have been taken to unearth the unique challenges to engage K-12 teachers and students in ML-empowered scientific discovery (SD) learning. Moreover, it is under-explored what scaffolding can be designed to mitigate the challenges. Based on our previous study and literature research, we identified three gaps for novice learners to conduct ML-empowered SD: (1) cognitive overload in ML visual analytics; (2) insufficient synthesis of multivariate patterns for hypothesis development; (3) the lack of evidence evaluation during the iterative investigation. We also propose three corresponding scaffolding components and evaluation studies for the next step.

CCS CONCEPTS

• **Human-centered computing → Activity centered design**; *User interface toolkits.*

KEYWORDS

Scaffolding Design, Scientific Discovery Learning, Machine Learning, K-12 Education

ACM Reference Format:
Xiaofei Zhou, Kaixin Li, Abdul Moid Munawar, and Zhen Bai. 2021. Scaffolding Design to Bridge the Gaps between Machine Learning and Scientific Discovery for K-12 STEM Education. In *Interaction Design and Children (IDC '21), June 24–30, 2021, Athens, Greece.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3459990.3465194

1 INTRODUCTION

Machine Learning (ML) is a promising tool for K-12 scientific discovery (SD) learning because it can reveal complex patterns to discover new knowledge from a large amount of data [19, 23] and the young generation needs to be prepared for the increasingly data-driven intelligent world [25, 27]. More research efforts have been made to introduce ML to K-12 students [9, 14, 17, 18]. However, little is known about how to design effective ML-empowered SD for K-12 students and teachers. The scientific explanation with ML-discovered results requires interpretational skills that are more challenging for young learners, such as making sense of anomalies detected and evaluating different data patterns by comparisons [13]. Beyond that, it would be even more demanding for K-12 students to steer the ML-empowered investigation (e.g., conducting ML experimentation, keeping track of multivariate hypotheses) [10, 22] while applying ML, the unfamiliar discovery tool, in relatively open-ended discovery learning.

To investigate applying ML as an SD tool for K-12 STEM education, we had designed a web-based learning environment named SmileyDiscovery [35], which makes k-means clustering accessible to novice learners in SD learning. Findings from our previous study with 18 K-12 STEM teachers [35] and further literature research indicate three major gaps for novice learners to interpret ML-discovered results during SD learning: (1) the extraneous load caused by translation between data and visual representation in ML analysis; (2) insufficiently synthesizing the ML-discovered evidence for hypothesis development; (3) cognitive challenges for young learners to evaluate different evidence forms to confirm/challenge hypotheses. To bridge these gaps, in this paper, we propose three scaffolding components informed by existing scaffolding design guidelines [22] and skills needed for scientific explanation [13], including (1) automatically parsing the graphical patterns in data visualization into patterns in data attributes, (2) facilitating hypothesis tracking and construction in multidimensional feature space during the iterative investigation, (3) hints adaptive to data patterns for learners to interpret. For the next step, we will evaluate the effectiveness of the new design components on facilitating K-12 students to interpret complex patterns revealed by ML, construct hypotheses progressively, and develop challenging scientific skills in evidence-based discovery. Our scaffolding components can engage young learners in a more meaningful and authentic ML-empowered SD learning experience beyond SmileyDiscovery. The initial design space proposed also aims to set the stage for a wide variety of ML-empowered learning environments in the future.

IDC '21, June 24–30, 2021, Athens, Greece

ACM ISBN 978-1-4503-8452-0/21/06.
https://doi.org/10.1145/3459990.3465194

2 RELATED WORK

2.1 Scaffolding Design for Computer-Supported SD Learning

Investigating meaningful and authentic science inquiries in different technology-enhanced learning environments [6, 24, 33] can empower learners to develop science knowledge and computational skills [10, 22]. To support learners' navigation in challenging and open-ended problems, a scaffolding design framework for Computer-Supported SD learning [22] has been developed for three constituent processes for inquiry: sense-making, process management, articulation & reflection. Recently, a few efforts have focused on how AI/ML can transform K-12 students' STEM learning [16, 18, 25, 31, 34–36]. However, existing scaffolding mainly involves technical facilitation or prompts provided by researchers manually [25, 34, 36] and data visualization that makes ML concepts or ML-generated results accessible [16, 31, 35, 36]. Thus, it is still an under-explored design space on what automated scaffolding can be designed to tackle the unique challenges in SD learning enhanced by ML for K-12.

2.2 Visual Analytics for ML

Many visual analytics systems have been developed to make sense of large and complex data processed by ML and other analytic techniques [4]. Pirolli and Card's sense-making loop model [21] illustrates how people gain insights from data by using visual analytic tools through a series of cognitive stages, including searching and storing relevant information, structuring the information for further explanation, constructing and testing hypotheses, and deriving conclusions. The Data-Frame Model [12] encourages people to question and reframe their understanding of data, which is the cognitive challenge for young learners to interpret different forms of evidence [13]. From the perspective of constructing visual representations for source data, the information visualization pipeline [8, 28] depicts the process of data transformation, visual mappings, and view transformation. Though being informative to facilitate learners' process management in science inquiry [22], existing visual analytics systems are not yet accessible for K-12 students who are still developing cognitive skills for scientific investigation. Thus, a corresponding scaffolding design is needed.

3 DESIGN SPACE FOR SCAFFOLDING ML-EMPOWERED SD LEARNING

3.1 Prior Work

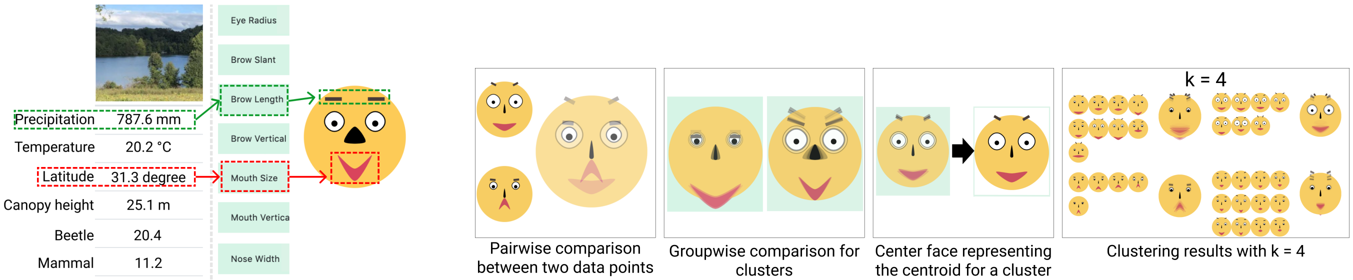
In our prior study, 18 K-12 STEM teachers applied ML in an SD learning activity about ecosystems using a web-based platform named SmileyDiscovery [35]. SmileyDiscovery utilizes glyph-based visualization [1, 32] to make multidimensional data accessible and playful for K-12 students by mapping individual data attributes (e.g., latitude) to graphical features of a face glyph [2] (e.g., mouth size) (Fig. 1a [35]). Besides, different glyph placements are provided for pattern interpretation (Fig. 1b [35]). We evaluated SmileyDiscovery's effectiveness of making cluster analysis interpretable for novices and its pedagogical potential in K-12 STEM learning by inviting teachers to interact and design their learning activities with SmileyDiscovery [35]. Furthermore, we identified the gaps teachers encountered while conducting SD with SmileyDiscovery by analyzing their discovery process with system log data and their feedback on the learning experience [35]. In Section 3.2 and Section 3.3, an overview of the scaffolding design space and corresponding gaps are discussed, respectively.

3.2 Scaffolding Design Space for ML-Empowered SD Learning

Based on the empirical evidence collected and further literature research, we propose an initial scaffolding design space for K-12 ML-empowered SD learning (Fig. 2). It focuses on the flow of data interpretation for hypothesis generation and evolvement in ML-empowered SD. Fig. 2(1), informed by the information visualization pipeline [33], shows a necessary step of transforming complex multidimensional data and ML-discovered results into accessible visual representations for K-12 students to observe. Fig. 2(2), similar to the "read&extract" step in the sense-making loop [21], refers to the cognitive stage where learners interpret visual representations into data patterns as evidence for further scientific explanation and hypothesis evolvement. For example, an analyst observes a cluster plot and interprets intra-cluster patterns for further synthesis and explanation. Fig. 2(3) targets confirming/challenging the hypothesis by interpreting patterns discovered. It points out an essential cognitive skill for scientific discourse: evaluating different evidence forms based on the claims derived [13], in addition to the CER (reasoning how the evidence supports the claim) procedure [7, 15] that is commonly used to facilitate scientific explanation in K-12 classrooms.

3.3 Identified Gaps for New Scaffolding Components

We identified three gaps in the design space that need new scaffolding components. *First*, ML visual analytics [33] requires learners to connect visual forms to source data for sense-making in the science context. This brings extra cognitive load to SD, which already has various memory requirements [11], impairing learners' performance in extracting evidence from data patterns to investigate the target science inquiry (Fig. 2 *G#1*). In our previous study, learners reported that it is overwhelming to translate patterns in graphical glyph features (e.g., big mouth and small eyes) back to science context with data attributes (e.g., high temperature and small beetle richness) [35]. *Second*, from log data of teachers' text input and clicking behaviors [35], we observed that it is challenging for them to sufficiently synthesize all the ML-discovered patterns and evolve hypotheses accordingly (Fig. 2 *G#2*) due to their unfamiliarity with multivariate analysis [5, 26]. Though K-12 teachers were able to interpret multivariate patterns during the investigation (e.g., one cluster of field sites shares patterns of high temperature&precipitation&canopy, low latitude, and high beetle richness while another cluster has high temperatures, low precipitation&canopy&latitude, and low beetle richness), the hypotheses they derived tended to stay binary (e.g., precipitation and beetle richness are positively correlated) and lack adequate synthesis of the evidence collected. *Third*, novice learners have little experience interpreting anomaly cases or complex and seemingly controversial patterns revealed by ML from messy real-world data (Fig. 2



(a) Visual mapping for a face glyph.

(b) Glyph placements and interaction for more accessible data patterns revealed by ML.

Figure 1: Glyph visualization and interaction designed in SmileyDiscovery for the prior study.

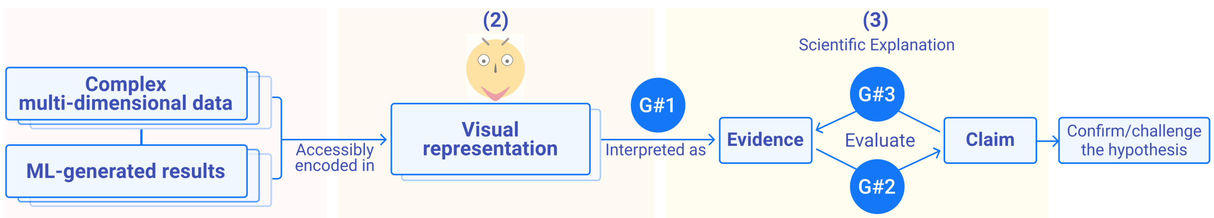


Figure 2: Design space and gaps identified for scaffolding in ML-empowered SD learning: *G#1.* Interpret visual representations into data patterns to collect evidence during the investigation; *G#2.* Iteratively derive claims by synthesizing multivariate evidence to confirm/challenge the hypothesis; *G#3* Evaluate different forms of evidence in relation to the claims.

G#3). There is a cognitive gap for young learners to distinguish the strengths and limitations for different evidence forms to confirm/challenge hypotheses [13]. For example, novice learners might find it confusing to confirm a hypothesis based on quantitative evidence from ML-discovered patterns when they observed outliers that seem to be counterexamples [35].

4 SCAFFOLDING COMPONENTS FOR K-12 ML-EMPOWERED SD LEARNING

4.1 S#1: Automatic Translation for Data Interpretation with Visual Representation

To engage learners more directly in the science context after observing visual representations and reduce the extra layer of cognitive load (Fig. 2 *G#1*), we propose the design to automatically parse learners' direct analysis of graphical glyph features into patterns in data attributes (Fig. 3). When the cursor hovers on a glyph element, a window will pop up containing corresponding glyph features. For example, hovering the cursor over the mouth of a face glyph activates the graphical features related to the mouth (i.e., the vertical location of the mouth, the size of the mouth) and sliders displaying the attribute values (Fig. 3(1)). While learners are dragging the thumb along a slider track, the corresponding graphical glyph feature changes to indicate how the attribute value influences the visual representation in real-time. Whenever learners stop interacting with a slider, the face and the slider reset the inputs to their original shape and value, respectively.

Taking such frequent and nonsalient tasks off students' cognitive load can better engage them in scientific inquiry with a deep cognitive focus on the salient activities [3, 22], such as how the ML-discovered patterns in data attributes can be interpreted and confirm/challenge the hypothesis. Furthermore, both ends of the slider indicate the minimum and maximum value of that attribute and thus contextualize individual data points in the entire dataset. This guides learners to step back and examine the larger context where they can interpret the values of data attributes more subjectively.

4.2 S#2: Develop Hypotheses Based on Multivariate Patterns Revealed by ML

To facilitate novice learners to synthesize multivariate patterns sufficiently (Fig. 2 *G#2*), we propose a new design for updating hypotheses on corresponding data attributes during the iterative scientific investigation (Fig. 4). Learners can construct hypotheses based on the multivariate patterns revealed by ML in a collapsible modal window (Fig. 4(1)-(2)). Hypothesis tracking on demand (Fig. 4(3)) restricts the complexity of hypothesis evolvement by setting informative boundaries, thereby enhancing learners' concentration during the investigation [22].

In conjunction with open-ended inquiry recording (Fig. 4(5a)-(5b)), the construction of a concept map (Fig. 4(4a)-(4c)) that encodes the direction and the magnitude of the correlations among multiple variables scaffolds young learners in an active meta-cognitive process to articulate multivariate correlations in scientific explanation step by step [22]. Such refined development can provide us with considerable specificity on what the learner already unearths that can be built upon, anchored in the zone of proximal

Figure 3: Automatic translation from (1) visual representation to (4) patterns in data attributes for further scientific explanation.

development [20, 29]. For example, after intra-cluster pattern interpretation reveals that one cluster of field sites shares high temperatures&precipitation&canopy, low latitude, and high beetle richness while another cluster has patterns of high temperatures, low precipitation&canopy&latitude, and low beetle richness, learners may infer that beetle abundance requires lush and humid environments. Thus, they can mark potential correlations in the hypothesis builder one by one (Fig. 4(4c)) for further validation with more ML-discovered evidence.

4.3 S#3: Hint Design to Evaluate Evidence for Iterative Investigation

To mitigate the cognitive challenges for young learners to evaluate different evidence forms (Fig. 2 *G#3*), hints to guide learners' pattern interpretation are designed as the third scaffolding component (Fig. 5). Based on different forms of patterns a learner could be currently interpreting [13], hints will be generated right beside to facilitate the learner to evaluate the scientific explanation and plan further investigation accordingly (Fig. 5a). For example, in cluster analysis, intra-cluster patterns revealed by a clustering algorithm suggest that mammal abundance requires warm & humid & lush environments. However, an outlier with low temperatures&precipitation&canopy but relatively high mammal richness may make learners hesitant about confirming the hypothesis. In this case, a hint can be provided to scaffold novice learners to evaluate the outlier and explore potential directions for further investigation. Fig. 5b shows an example designed for SmileyDiscovery.

Such hints adaptive to the paths of investigating a science inquiry [30] can provide structure for complex tasks in SD and enhance learners' ongoing reflection by highlighting the epistemic features of investigation [22]. Specifically, such hint design can address the cognitive challenge for young learners to establish a more critical and dynamic perspective on the evidence-claim relationship [13] and develop corresponding scientific skills. For data patterns that are not expected in the diagram (Fig. 5a), generic hints will be provided close at hand to inspire learners' next steps, such as "How could this result have happened?", "If my explanation is true, what patterns should also be found in the data?", "What are any other potential mechanisms that would have given these results?".

5 PILOT TEST AND FUTURE WORK

In our next phase, we will first pilot test the usability of each scaffolding component proposed above by implementing them in the next version of SmileyDiscovery. For *S#1*, the evaluation will focus on the efficiency and accuracy of learners completing pattern interpretation tasks compared to the condition of using the SmileyDiscovery from the prior study [35]. For *S#2*, the primary goal is to evaluate if it can help young learners synthesize complex ML-discovered patterns and evolve hypotheses more effectively. For *S#3*, we plan to evaluate the learnability of different hints, looking at how efficient they are for K-12 students to evaluate the evidence in relation to the claims.

After the pilot test, we will conduct both a teacher study and a student study. For the teacher study, we aim to gain a deeper understanding of the connections between ML-discovered patterns and SD practices in K-12 STEM education. For the student study, we plan to evaluate if ML-empowered SD learning can be beneficial for young learners to develop challenging scientific skills [13]. We also would like to investigate the effectiveness of the scaffolding design for other ML methods to empower K-12 SD learning.

REFERENCES

[1] Rita Borgo, Johannes Kehrer, David HS Chung, Eamonn Maguire, Robert S Laramee, Helwig Hauser, Matthew Ward, and Min Chen. 2013. Glyph-based Visualization: Foundations, Design Guidelines, Techniques and Applications.. In *Eurographics (STARs)*. 39–63.
[2] Herman Chernoff. 1973. The use of faces to represent points in k-dimensional space graphically. *Journal of the American statistical Association* 68, 342 (1973), 361–368.
[3] Mihaly Csikszentmihalyi and Mihaly Csikzentmihaly. 1990. *Flow: The psychology of optimal experience*. Vol. 1990. Harper & Row New York.
[4] Alex Endert, William Ribarsky, Cagatay Turkay, BL William Wong, Ian Nabney, I Díaz Blanco, and Fabrice Rossi. 2017. The state of the art in integrating machine learning into visual analytics. In *Computer Graphics Forum*, Vol. 36. Wiley Online Library, 458–486.
[5] National Governors Association Center for Best Practices. 2010. C. o. C. S. S. O. Common Core Science Standards. National Governors Association Center for Best Practices, Council of Chief State School.
[6] Janice D Gobert, Michael Sao Pedro, Juelaila Raziuddin, and Ryan S Baker. 2013. From log files to assessment metrics: Measuring students' science inquiry skills using educational data mining. *Journal of the Learning Sciences* 22, 4 (2013), 521–563.
[7] Amelia Wenk Gotwals and Nancy Butler Songer. 2010. Reasoning up and down a food chain: Using an assessment framework to investigate students' middle knowledge. *Science Education* 94, 2 (2010), 259–281.
[8] Jeffrey Heer, Stuart K Card, and James A Landay. 2005. Prefuse: a toolkit for interactive information visualization. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 421–430.
[9] Ken Kahn and Niall Winters. 2017. Child-friendly programming interfaces to AI cloud services. In *European Conference on Technology Enhanced Learning*. Springer, 566–570.
[10] Tanja Käser and Daniel L Schwartz. 2020. Modeling and Analyzing Inquiry Strategies in Open-Ended Learning Environments. *International Journal of Artificial Intelligence in Education* 30, 3 (2020), 504–535.
[11] David Klahr and Kevin Dunbar. 1988. Dual space search during scientific reasoning. *Cognitive science* 12, 1 (1988), 1–48.
[12] Gary Klein, Brian Moon, and Robert R Hoffman. 2006. Making sense of sensemaking 2: A macrocognitive model. *IEEE Intelligent systems* 21, 5 (2006), 88–92.
[13] Deanna Kuhn and Daniel Lerman. 2021. Yes but: developing a critical stance toward evidence. *International Journal of Science Education* (2021), 1–18.

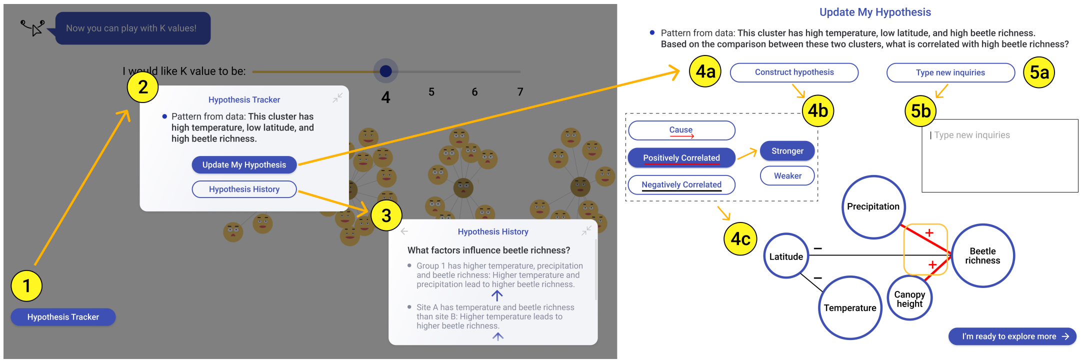
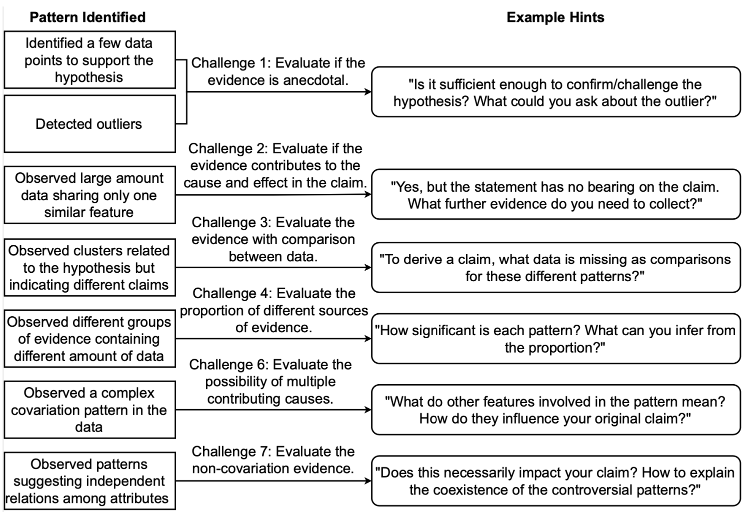
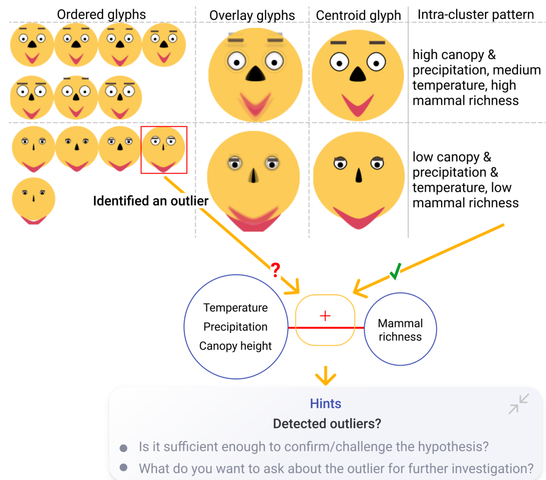


Figure 4: Update the hypothesis by synthesizing the multivariate patterns revealed by ML.



(a) Hints adapted for ML-discovered patterns identified by learners.



(b) An example of hint generation when learners observe an outlier in SmileyDiscovery.

Figure 5: Hint design to evaluate different forms of evidence during the iterative investigation

[14] Irene Lee, Safinah Ali, Helen Zhang, Daniella DiPaola, and Cynthia Breazeal. 2021. Developing Middle School Students' AI Literacy. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education*. 191–197.

[15] Haiying Li, Janice Gobert, and Rachel Dickler. 2017. Automated Assessment for Scientific Explanations in On-Line Science Inquiry. *International Educational Data Mining Society* (2017).

[16] Phoebe Lin, Jessica Van Brummelen, Galit Lukin, Randi Williams, and Cynthia Breazeal. 2020. Zhorai: Designing a Conversational Agent for Children to Explore Machine Learning Concepts. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 34. 13381–13388.

[17] Duri Long and Brian Magerko. 2020. What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–16.

[18] Lívia S Marques, Christiane Gresse von Wangenheim, and Jean CR HAUCK. 2020. Teaching Machine Learning in School: A Systematic Mapping of the State of the Art. *Informatics in Education* 19, 2 (2020), 283–321.

[19] Michael Muller, Shion Guha, Eric PS Baumer, David Mimno, and N Sadat Shami. 2016. Machine learning and grounded theory method: convergence, divergence, and combination. In *Proceedings of the 19th international conference on supporting group work*. 3–8.

[20] Joseph D Novak. 2005. Results and implications of a 12-year longitudinal study of science concept learning. *Research in Science Education* 35, 1 (2005), 23–40.

[21] Peter Pirolli and Stuart Card. 2005. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proceedings of international conference on intelligence analysis*, Vol. 5. McLean, VA, USA, 2–4.

[22] Chris Quintana, Brian J Reiser, Elizabeth A Davis, Joseph Krajcik, Eric Fretz, Ravit Golan Duncan, Eleni Kyza, Daniel Edelson, and Elliot Soloway. 2004. A scaffolding design framework for software to support science inquiry. *The journal of the learning sciences* 13, 3 (2004), 337–386.

[23] Charles Romesburg. 2004. *Cluster analysis for researchers*. Lulu.

[24] Jonathan Rowe, Bradford Mott, Scott McQuiggan, Jennifer Robison, Sunyoung Lee, and James Lester. 2009. Crystal island: A narrative-centered learning environment for eighth grade microbiology. In *workshop on intelligent educational games at the 14th international conference on artificial intelligence in education, Brighton, UK*. 11–20.

[25] Shashank Srikant and Varun Aggarwal. 2017. Introducing data science to school kids. In *Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education*. 561–566.

[26] Ngss Lead States. 2013. *Next Generation Science Standards: For States, By States*. The National Academies Press, Washington, DC.

[27] David Touretzky, Christina Gardner-McCune, Fred Martin, and Deborah Seehorn. 2019. Envisioning AI for k-12: What should every child know about AI?. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 33. 9795–9799.

[28] Jarke J Van Wijk. 2005. The value of visualization. In *VIS 05. IEEE Visualization, 2005*. IEEE, 79–86.

[29] Lev Semenovich Vygotsky. 1980. *Mind in society: The development of higher psychological processes*. Harvard university press.

[30] Maaike Waalkens, Vincent Aleven, and Niels Taatgen. 2013. Does supporting multiple student strategies lead to greater learning and motivation? Investigating a source of complexity in the architecture of intelligent tutoring systems. *Computers & Education* 60, 1 (2013), 159–171.

[31] Xiaoyu Wan, Xiaofei Zhou, Zaiqiao Ye, Chase K Mortensen, and Zhen Bai. 2020. SmileyCluster: supporting accessible machine learning in K-12 scientific discovery. In *Proceedings of the Interaction Design and Children Conference*. 23–35.
[32] Colin Ware. 2019. *Information visualization: perception for design.* Morgan Kaufmann.
[33] Carl E Wieman, Wendy K Adams, and Katherine K Perkins. 2008. PhET: Simulations that enhance learning. *Science* 322, 5902 (2008), 682–683.
[34] Yuanlin Zhang, Jianlan Wang, Fox Bolduc, William G Murray, and Wendy Staffen. 2019. A preliminary report of integrating science and computing teaching using logic programming. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 33. 9737–9744.
[35] Xiaofei Zhou, Jingwan Tang, Daley Michael, Saad Ahmad, and Zhen Bai. 2021. "Now, I Want to Teach it for Real!": Introducing Machine Learning as a Scientific Discovery Tool for K-12 Teachers. (2021). (In press).
[36] Abigail Zimmermann-Niefield, Makenna Turner, Bridget Murphy, Shaun K Kane, and R Benjamin Shapiro. 2019. Youth learning machine learning through building models of athletic moves. In *Proceedings of the 18th ACM International Conference on Interaction Design and Children*. 121–132.