DPV (Domain, Purpose, Visual) Framework: A data visualization design pedagogical method for middle schoolers

Yixuan Li[1], Alex Endert[1] and Jessica Roberts[1]

[1] Georgia Institute of Technology, USA

Figure 1: *Samples of Students' final artifacts in the showcase*

Abstract

Data visualization literacy is essential for K-12 students, yet existing practices emphasize interpreting pre-made visualizations rather than creating them. To address this, we developed the DPV (Domain, Purpose, Visual) framework, which guides middle school students through the visualization design process. The framework simplifies design into three stages: understanding the problem domain, specifying the communication purpose, and translating data into effective visuals. Implemented in a two-week summer camp as a usage scenario, the DPV framework enabled students to create visualizations addressing community issues. Evaluation of student artifacts, focus group interviews, and surveys demonstrated its effectiveness in enhancing students' design skills and understanding of visualization concepts. This work highlights the DPV framework's potential to foster data visualization literacy for K-12 education and broaden participation in the data visualization community.

CCS Concepts

• ***Human-centered computing*** → ***Visualization design and evaluation methods;*** *Empirical studies in visualization;*

1. Introduction

Data Visualization Literacy, the ability to read and construct data visualizations, has been addressed in 21st-century [BBG19]. As a result, Data Visualization Literacy education is also increasingly addressed in K-12 scenarios [BMBH16]. For instance, in K-12 education standards in the US, students are required to understand science concepts through charts and create data visualizations. However, while these skills have been widely addressed in educational research [BBG19,IFMTW23], the complexities of the visualization design process are often overlooked in K-12 practices.

In higher education, visualization researchers have developed robust pedagogical methodologies, including open-access curricula, worksheets, and interactive tools [HBH*20, KHP*23, BKR*23]. Examples include the design activity worksheets for graduate students to bridge theory and practice [MLM17], and the DVA worksheet based on visualization process model [BYP21]. Physical materials are designed to engage students through hands-on activities [HA17, BSK*24]. However, these methods are designed and tested primarily for graduate or professional students for higher education, leaving their effectiveness in K-12 settings unclear. Limited tools tested with younger students, such as the Comic Construction Kit, indicate that additional scaffolding is needed [BSK*24]. This may stem from challenges unique to K-12 learners, such as limited prior knowledge, the need for accessible resources, and reliance on scaffolded guidance [BYP21, Shr18]. These challenges underscore the need to adapt existing frameworks—or create new ones—tailored to K-12 learners and educators.

To address this gap, we developed the DPV (Domain, Purpose, Visual) framework, integrating insights from K-12 teachers and higher education methodologies. This framework guides K-12 stu-

DOI: 10.2312/eved.20251023

dents through data visualization design in three stages: understanding the problem domain, specifying the communication purpose, and translating data into effective visuals. We employed the DPV framework in a two-week summer camp for middle schoolers in 2024. A mixed-method analysis of student work and focus groups showed its effectiveness in helping students create visualizations to communicate their ideas. Student feedback was also collected to inform ongoing refinements, making it adaptable to the needs of K-12 learners. This paper provides a detailed account of adapting data visualization education for K-12 students. By explaining the DPV framework's development, implementation, and impact, this study provides resources for educators and researchers to empower K-12 students with visualization skills and promote diversity and inclusiveness in the Vis community.

2. Related Work

Our work is inspired by previous efforts in K-12 data visualization education, design frameworks, and pedagogical frameworks in higher education, as described in the following sections.

2.1. Data Visualization Education for K-12

Data visualization is a core aspect of data science education, yet most K-12 curricula focus on interpreting pre-made visualizations rather than creating them [BMBH16]. For example, students analyze visualizations to uncover biases [POPR16], evaluate data quality [Dea14], or reflect on self-collected data through statistical charts [LDCT21]. While these approaches build analytical skills, they limit opportunities for students to engage in the creative processes of visualization design. Several studies highlight the need for more data visualization activities in K-12, enabling students to interpret graphical displays and create their own representations to uncover patterns and communicate ideas. For instance, the Maine Data Literacy Project offers various teaching materials for middle and high school teachers to learn data and graphs in school setups [WGM*16]. PlayData, a block-based programming environment for data visualization purposes, has been used in multiple camps targeting data science and data visualization [FBL24]. These practices often emphasize high-level tasks like questioning, inference, and storytelling, while neglecting essential design skills such as selecting representations, mapping data to encodings, optimizing designs for their purpose, and evaluating design impacts.

Addressing this gap, the DPV framework aims to guide students through the data visualization design process, from high-level developing research questions and communicating ideas to low-level mapping data to visual encodings, fostering a deeper understanding of how design decisions influence data communication.

2.2. Design Frameworks in Data Visualization

Researchers in the data visualization community have developed structured approaches to enhance designers' workflows. The Four-Level Nested Model [Mun09], proposed by Munzner, structures information visualization (infovis) design into four levels: (1) domain problem characterization, (2) data/operation abstraction design, (3) encoding/interaction technique design, and (4) algorithm design. Sedlmair et al. [SMM12] expanded [Mun09] with a nine-stage design study framework, divided into precondition, core, and analysis phases. This framework emphasizes iterative design, problem characterization, implementation, and reflective analysis in the design process. McKenna et al. [MMAM14] introduced the Design Activity Framework, aligning iterative activities—understand, ideate, make, and deploy—with [Mun09]. It connects design actions to decision-making at each nested level. Meyer et al. [MSQM15] extended the model with the Nested Blocks and Guidelines Model (NBGM), adding "blocks" to represent design outcomes and "guidelines" to link decisions across levels, improving its applicability to complex visualization tasks. Roberts et al.'s Five Design-Sheet (FdS) Methodology [RHR15] emphasizes sketching for divergent thinking and iterative refinement, aligning with the first three levels of [Mun09] to guide early-stage planning and prototyping of visualizations.

Existing frameworks emphasize contextualizing design through domain understanding and fostering creativity by generating multiple design ideas. However, derived by DataVis experts, these frameworks are too complex and heavy for K-12. Our work develops a pedagogical framework tailored for K-12 learners, enabling them to engage in design processes akin to those of professionals.

2.3. Pedagogical Practices for Data Visualization Design

Data visualization is a well-established research area with courses widely offered in higher education for undergraduate and graduate students across disciplines. Researchers have created pedagogical methods [BYP21, RRJH17] for teaching data visualization, including open-access curricula, worksheets, card kits, and comic-making tools. For example, McKenna et al. [MLM17] developed design activity worksheets based on [MMAM14]. These worksheets offer a structured guide for students, connecting theoretical frameworks with practical teaching. They were evaluated with 66 students in a graduate-level course at the University of Utah. Another study [Fry07] adapted Fry's seven-stage visualization process into data visualization activity (DVA) worksheets to assist teachers new to data visualization. These worksheets were tested in a CS undergraduate course at Purdue University Researchers also explored physical materials to guide students through the design process. VIZITCARDS [HA17] is a card-based workshop for the graduate-level information visualization course at the University of Michigan, designed to promote engagement and collaboration among students. Comic-based toolkits are also regarded as effective pedagogical tools for teaching data visualization to a broader audience. The Comic Construction Kit [BSK*24] is designed to help students learn different chart types and techniques by creating explanatory comics in a workshop format. Testing revealed that pre-college students needed more instructional support than college, grad students, and faculty, indicating that additional resources are necessary for K-12 educational settings.

These works establish a valuable foundation for teaching data visualization, revealing that worksheet-based methods effectively support the iterative design process by breaking complex steps into manageable activities. Our work builds on this by customizing structured design processes and scaffolding to better meet K-12 learners' needs for accessible terminology and hands-on guidance.

3. Design Goals and Iterations

Our goal is to design a pedagogical framework to teach K-12 students data visualization in a summer camp setup. The framework addresses challenges by incorporating educational standards, teacher feedback, and an iterative development process.

3.1. Requirements from K-12 Education

The design requirements were identified from the data visualization skills students need per educational standards and the scaffolding teachers require based on a needs assessment survey.

3.1.1. Learning Objectives

To ensure our pedagogical framework is practical and applicable, it aligns with K-12 education standards. In the United States, teachers are required to align their teaching to pre-designed national or state standards in math, science, and language arts. The standards are indexed to provide alignment across grades. For example, physical science standards on energy begin with the grade level (e.g. 4 for fourth grade), then the area (PS for physical science), followed by a number (e.g. standards about energy are numbered 3-1, 3-2, etc.). So a 4th-grade student would be expected to master 4-PS-3-1: *Use evidence to construct an explanation relating the speed to the energy of that object.* By middle school, data visualization skills are explicitly integrated. For instance, MS-PS3-1 (the middle school equivalent of the standard above) dictates students should *construct and interpret graphical displays of data to identify relationships of kinetic energy to the mass of an object and to the speed of an object.* Similarly, the Common Core English Language Arts Standards for Reading in Science and Technical Subjects for middle school grades 6-8 (abbreviated CCSS.ELA-Literacy.RST6-8) asks students to *translate quantitative or technical information expressed in words in a text into visual form and translate information expressed visually or mathematically into words*. These standards demonstrate the need for students to develop skills in interpreting, creating, and critically evaluating data visualizations. Specifically, students are required to learn to interpret information from visualizations, create visualizations to support their scientific communication, and critically assess the effectiveness of their designs.

Building on these requirements, we derived the following learning objectives the framework aims to accomplish:

- **LO1:** Analyze a given dataset to recognize patterns, make inferences, draw conclusions, and develop a narrative.
- **LO2:** Identify the strengths and weaknesses of a visualization design for effective data communication.
- **LO3:** Understand and apply the data visualization design process to create visualizations for communicating ideas.
- **LO4:** Use digital tools to create data visualizations.

3.1.2. K-12 Teachers' Needs

It is also important to understand what the current practitioners need for the pedagogical framework we develop. We sent out a teachers' needs assessment survey with IRB approval [†] in Fall 2023 through an email list of a local science festival, reaching science teachers mainly in the local city but also in the US, to understand the current practice and challenges that teachers meet when teaching with data visualization. 11 participants consented to participate, with five of them middle school teachers, three high school teachers, two K-5 teachers, and one higher education teacher.

The survey included four multiple-selection questions with text input to elaborate on their choices, covering chart types used, chart sourcing methods, and experiences and challenges in teaching with data visualizations. We analyzed teachers' selection responses to those four questions. While all participants reported using data visualizations in their classes, and 7 teachers generally had positive experiences, 9 teachers noted facing challenges. The most common issues were insufficient resources (5 responses) and students' difficulties in understanding charts (5 responses). Other challenges included time constraints (3 responses) and limited student interest (3 responses). These findings show a need for accessible resources and materials that can provide structured guidance for educators. Time constraints and students' varying levels of interest and comprehension further underscore the importance of practical and adaptable strategies for K-12 educators. Incorporating these needs into the design of the DPV framework, we aimed to empower teachers with resources and methods to enhance their teaching practices, making data visualization more accessible and engaging for their students.

3.2. Design Goals

We established key design goals from these requirements to ensure relevance in K-12 education. These goals address challenges from our survey and align with our **LO**s, creating a framework that supports skill development and meaningfully engages students.

- **DG1: Enhance Data Visualization Skills** The framework aims to equip students with the ability to interpret existing data visualizations and design their own, fostering essential skills for both academic and real-world applications.
- **DG2: Accessible for Middle Schoolers** To ensure the usability of the framework for our target K-12 learners, the framework should be age-appropriate, intuitive, and adaptable to the developmental levels and learning needs of middle school students.
- **DG3: Foster Creativity and Critical Thinking** The DPV framework should encourage students to consider multiple visualization ideas, critically evaluate them, and refine their designs, promoting creative exploration and analytical thinking.
- **DG4: Provide Authentic Experience** The framework should scaffold students to progress through an authentic design experience consisting of high-level considerations (such as contexts and audience) and low-level visualization design considerations.

In the following section, we outline the process of developing the DPV framework to achieve these design goals. We later implemented the framework during a two-week summer camp and evaluated its effectiveness in meeting the design goals outlined above.

3.3. Iterations and Refinement

We outline the iterative process of developing and refining the DPV framework. The adjustments of each version are informed by K-12 teachers to ensure the framework is aligned with their needs.

[†] IRB Protocol Number: H23430

The Four-level Nested Model [Mun09]	The DPV Framework
Domain problem characterization	**Domain** · What data are you going to use? · Who is your audience? · What are domain conventions?
Data/operation abstraction design	**Purpose** · What information is shown? · What are users' tasks? · How do you want to use this vis?
Encoding/interaction technique design	**Visual** · Chart Types · Axes · Marks · Channels
Algorithm design	Beyond the learning objectives

Table 1: *Comparing the Four-level Nested Model [Mun09] with the first version of the DPV Framework*

3.3.1. Version 1: Initial framework design

The Four-Level Nested Model [Mun09] provides a design foundation considering **DG1** and **DG4**. It progresses from domain problem characterization to encoding design, linking high-level tasks with low-level execution. We started by making adaptations to this framework considering our design goals, as shown in Table 1. Terminology was adjusted to enhance accessibility, with "Abstraction Design" renamed to Purpose and "Encoding" renamed to Visual (**DG3**). Guiding sub-questions were included under each stage to scaffold students' design actions (**DG2**). The framework concentrated on the first three levels considering the **LO**s. Additionally, a final step for creating a sketch was incorporated to help students connect their conceptual understanding with practical application. Unlike Munzner's [Mun09] or Sedlmair's [SMM12] frameworks, which are often used by expert designers, the DPV framework's items are simplified and education-oriented, designed to prompt concrete, actionable, and age-appropriate reflections from students rather than eliciting comprehensive analytical justifications.

3.3.2. Version 2: Feedback from teachers' workshop

Before testing with students, we recruited seven K-12 teachers through a 4-hour professional development program for a pilot workshop to test our DPV framework with IRB approval [‡]. We introduced the framework and demonstrated its application. Teachers then participated in a design activity using a dataset of air quality to create a chart to teach this topic. Finally, we held a 45-minute focus group with teachers who consented to participate to discuss their suggestions on using the DPV framework for instruction.

Through analyzing the transcript from the focus group using the thematic analysis method [BC06a], we identified several actionable suggestions for refining the DPV framework, primarily focused on **DG2** and **DG4**. For example, teachers suggested drawing from students' prior knowledge at the Domain stage to help them better understand their data and articulate their problems clearly. Also, using age-appropriate language was recommended by teachers, such as using terms like "Big Idea" instead of "Goal" was recommended to make it more approachable for students. At the Visual stage, teachers suggested providing more explicit options like "color, size, shape" instead of terms like "visual channels," to clarify what students should be designing. Finally, they also suggested adapting the worksheet to support collaborative and project-based learning.

‡ IRB Protocol Number: H24269

3.3.3. Version 3: Refine with a local teacher

The research team collaborated with a middle school science teacher from the workshop to finalize the DPV framework worksheet. In addition to earlier suggestions, the teacher proposed adding a sub-question "Problem Statements" under the domain stage to encourage students to research their problem setup further. The teacher also emphasized the importance of students creating sketches after completing the DPV framework and refining their designs iteratively. Together, we developed activities, lessons, and worksheets centered on the DPV framework. The finalized worksheet, included in the supplemental materials, provides prompt questions along with space for writing and sketching to help students articulate and develop their ideas. It was distributed during a two-week data visualization summer camp where students used it to guide their projects, as detailed in the next section.

4. The DPV Framework

The DPV (Domain, Purpose, Visual) framework is designed to guide K-12 students through the process of creating data visualizations for communication. Considering the **DG**s, it breaks down the design process into three stages: **Domain**, which focuses on understanding the context; **Purpose**, which emphasizes specifying the communication goal; and **Visual**, which centers on translating data into visual representation. Each stage includes structured questions to scaffold students' decision-making. Through the DPV framework, students learn to investigate data questions, create effective visualizations, and evaluate their design choices.

4.1. Domain: Understand the problem

The Domain stage requires students to first develop a more in-depth understanding of the domain areas their data visualization wants to address. This includes articulation of the topic, the current situation of the problem they identify, the potential audiences of their visualizations, and the conventions in this domain.

- **What is your topic?** This question prompts students to define the focus of their design. It encourages them to clearly identify the central subject or theme they want to explore, such as air quality or graduation rates. By narrowing down a topic, students establish a concrete foundation for their visualization, ensuring that their work is meaningful to a specific area of inquiry.
- **What is your problem statement?** Here, students are asked to describe the current situation and challenges of their chosen topic. They should research existing work that helps refine key concerns and uncover best practices. This step enables students to frame their visualizations around real-world applications.
- **Who is your audience?** This question prompts students to identify their target audience, such as local policymakers, educators,

students, or community groups. Understanding the audience is essential for tailoring the design to their needs and preferences, making the visualization more impactful and meaningful.

- **What are the domain conventions?** Students are prompted to recognize the common practices and standards used in their domain of interest, including specific visual codes or color schemes. For instance, in the partisan divide, red represents Republicans, and blue represents Democrats. Misusing these color schemes can result in the misrepresentation of domain-specific issues and hinder interpretation.

Addressing these sub-questions helps students understand their visualization's context and purpose, laying the groundwork for effective and meaningful design decisions in their later design.

4.2. Purpose: Specify the Goal

The Purpose stage focuses on helping students define the intent and communication goal of their data visualization. This step bridges the understanding of the domain with the creation of visualizations by ensuring students have a clear purpose for their work.

- **What is the Big Idea type?** We adopt the chart-type decision tree by [Abe21] to guide students in aligning designs with communication goals. It prioritizes the purpose of the visualization before selecting a chart type, focusing on four "Big Ideas": (1) Comparison: Highlight differences, e.g., comparing nutritional values of school lunches and fast food. (2) Distribution: Show spread or clustering, e.g., mapping fast-food locations in a neighborhood. (3) Relationship: Reveal correlations, e.g., linking forest loss to housing development. (4) Composition: Depict parts of a whole, e.g., food waste source breakdowns. Breaking the purpose into four categories helps students understand the concept and explore different ways to represent data. By defining the "Big Idea," they clarify their visualization's intent, ensuring their design choices effectively support their communication goals.
- **How will you use this visualization to tell your story?** This question prompts students to critically consider how their visualization communicates their story—whether by providing background, explaining a situation, or proposing a solution. For example, a visualization to provide background might include an overview of historical data, setting the stage for the discussion. A chart to explain a situation could focus on identifying and analyzing key patterns or relationships in the data, helping the audience understand a specific issue. Reflecting on these different usages helps students design strategically aligned visualizations, enhancing their narrative's impact and engagement.

4.3. Visual: Connect Data with Visual

The Visual stage focuses on mapping data into visual elements, ensuring that students make intentional design choices that align with their story. By considering how data is mapped to visuals, students enhance both the clarity and effectiveness of their visualizations.

What visuals are you going to use to tell your story (color, size, position, etc)? This question prompts students to think about how they will connect their data to specific visual elements to effectively communicate their story. It encourages them to make deliberate choices about visual channels such as color, size, position, and shape, which are essential for encoding information, and compare different design ideas to optimize their communication. For instance, students might use color to represent categorical variables, such as different types of pollutants in an air quality dataset, or use size to highlight the magnitude of a variable, such as the volume of waste produced by different sources. Position, a fundamental visual channel, might be used to place data points along axes to show trends or relationships. This step challenges students to think critically about which visual elements best align with their data and story, ensuring clarity and impact in their final visualization. By making these design decisions, students also develop an understanding of how visual encodings can enhance or hinder interpretation, which is a crucial skill in data visualization.

4.4. Sketch It Up

Sketching is a critical, standalone step in the DPV framework. It allows students to translate their ideas into a tangible draft, experimenting with various layouts, chart types, and design elements before creating a digital visualization. This phase helps visualize how data will be represented and ensures that visual elements align with their intended narrative. Through sketching, students can identify challenges, explore alternatives, and refine their designs while receiving valuable feedback from peers or instructors. This hands-on approach actively engages students in the design process, bridging the gap between planning and implementation iteratively.

5. Camp Experience and Evaluations

In the summer of 2024, the research team worked together to co-design and co-lead a Data Visualization and Environment (DVE) summer camp. This two-week camp aims to prepare middle schoolers to create data visualization to investigate real-world data-related challenges and communicate their data insights. We implemented the DPV framework during the summer camp and evaluated the framework according to the design goals illustrated in Sec. 3.

23 middle school students (average age = 12.2, 13 identifying as female) were recruited from a local school. While 68% of students had prior experience using data in school projects, most had not participated in data-focused programs before. During the first week, students formed into groups and identified their project topics as they practiced data visualization skills through structured problems. For the second week, students worked with their groups on the projects following the DPV worksheet. On the last day of the camp, students presented their work in an exhibition space on campus, where parents and community stakeholders were invited.

In this section, we will describe the metrics we devised to interpret students' use of the DPV framework and their general experience of this data visualization summer camp.

5.1. Data

This IRB-approved study[§] was conducted with the consent/assent of parents/guardians, students, and teachers. Students completed a

§ IRB Protocol Number: H24298

survey assessing their skills growth throughout the summer camp. We conducted 4 30-45-minute focus groups of 5-6 campers with parental consent each, which were audio recorded and transcribed. We used students' assigned IDs from focus groups to indicate them in the following sections (such as FG3-1, meaning Camper 1 in Focus Group 3) regarding their comments from the focus group. We also collected students' worksheets and final projects. We reported this data by their project topic (such as *Gas Price* Team). We then conducted a mixed-method analysis of the study data.

5.2. Students' Final Projects

The 23 students formed 10 groups, each identifying a community-related issue for investigation. Through one week's effort on their projects following the DPV framework, all groups succeeded in producing a final work that broadly touched on different perspectives of their communities, such as grocery, education, and environment. Students created data visualizations using tools like Tableau and Google Sheets and then curated them into posters for the final showcase. Fig. 1 displays examples of students' projects.

We analyzed students' focus group responses about using the DPV framework in creating visualizations to see how it supported their design processes and improved their data visualization skills targeting **DG1**. When asked,"What is your topic, and who is your audience?" all students selected topics that demonstrate connections to their interests and experiences. This personal relevance motivated their engagement in their data visualization design. Students were able to identify their audiences for their visualizations, thoughtfully considering who might be curious about their topics or who the key stakeholders were, supporting **DG4**. For instance, the *School Lunch* Team (Fig. 1.A) focused on the "nutritional value of fast food and school lunch," targeting the school board, students, and cafeteria staff. Meanwhile, the *Graduation* Team (Fig. 1.B) explored "graduation rates in their county," aiming to reach educators, parents, and the Department of Education. Students also demonstrated the ability to make logical and effective design choices based on domain conventions. For instance, the *School Lunch* Team used red to represent Chick-fil-A, aligning with the brand's primary color, and blue for their school dining hall, corresponding to its main color. Similarly, the *Graduation* Team used the primary colors of each school in their charts for clarity and relevance.

When asked, "What Big Idea do you want your audience to take away?" students were able to articulate their key message clearly in their own words, supporting **DG1&4**. Their goals ranged from persuading stakeholders to address specific issues (e.g., "Tell the school principals that we don't like school food because of health concerns" by the *School Lunch* Team), to providing new information to interested audiences (e.g., "To figure out if we were eating the right thing in life" by the *Fast Food* Team), and supporting decision-making (e.g., "How to save animal habitats from land development" by the *Housing* Team, Fig. 1.D). Their final projects showcased various Big Idea Types: nine focused on comparison (e.g., the *Fast Food* Team comparing store numbers), three on relationship (e.g., the *Housing* Team linking forest loss to housing developments), two on composition (e.g., the *Food Waste* Team analyzing food waste sources, see Fig. 1.C), and one on a distribution (the *Fast Food* Team mapping fast food stores in their community).

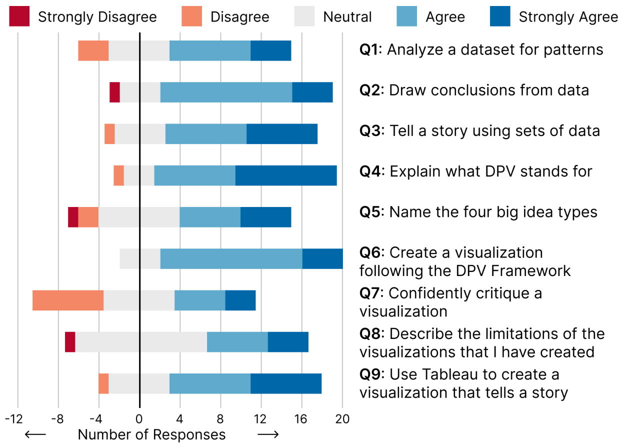

Figure 2: *Students' self-evaluation in the post-survey*

Students demonstrated use of a variety of chart types and visual channels in their final work across 10 projects, supporting **DG3**. Bar charts were the most common chart type, appearing 14 times in 9 projects. Real-world glyphs, like trash bags from the *Food Waste* Team and school mascots from the *Graduation* Team, add creativity and contextual relevance. Other chart types, including scatterplots, pie charts, and maps, were also used, reflecting students' ability to employ diverse design ideas to communicate their findings. Color was the most frequently used channel (15 times), often representing categorical variables like different fast food stores, counties, or schools. Size and position were also frequently utilized (12 and 6 times, respectively) to encode quantitative variables. Shape was used the least, appearing in only one project to represent categorical data. These results show that students made thoughtful design choices, effectively employing a variety of visual representations to create high-quality visualizations to support communication.

5.3. Students' experience with the DPV Framework

To understand students' experience of the DPV Framework, we evaluated the usability of the framework and a series of indicators of how students achieved the learning objectives of the camp. We discussed our investigations and observations of students' focus group transcripts, their final artifacts, and the post-survey.

Have students developed their data visualization skills (DG1)? We assessed students' self-efficacy regarding the learning objectives through a post-survey. Students rated their perceived skills on nine questions through a 5-point Likert scale (1 = Strongly Disagree to 5 = Strongly Agree). The results, presented in Fig. 2, showed generally positive self-efficacy in data literacy skills, with significantly more agreement than disagreement, particularly for LO1, LO2, and LO4. Statements about the DPV framework (Q4, Q5) and data communication (Q2, Q3, Q9) received ratings of 4 or 5 from over 18 students, indicating success in using the DPV framework to address community issues and high self-efficacy in applying data skills. However, for LO3, students were less confident critiquing a visualization (Q7), and many showed neutrality in critiquing their own (Q8). This indicates that while students can critique their visualizations to make design decisions, more guidance is needed for critiquing others' data visualizations.

We annotated the types of charts students initially intended to use, as indicated in their DPV worksheet, and the charts they used in their presentations. This approach helps us evaluate students' skills in choosing different data visualization methods to support their arguments. While students predominantly utilized bar charts in their DPV worksheet, they incorporated a broader range of visualization types in their final presentations, reflecting an increasing diversity in their choices. In the worksheet specifically, bar charts accounted for 46.7% of visualizations, making them the most common type. Other types accounted for about 6-14% of uses. In the Presentation stage, bar charts remained the most popular at 39%, although their use slightly declined. Meanwhile, alternatives like hand-drawn illustrations, videos, games, and 3D printing represented nearly 20% of the final visualizations. This change highlights the growing diversity and creativity in students' work, showing that our teaching approach encourages them to go beyond traditional formats and effectively share their ideas.

Is the DPV framework easy to understand (DG2)? In the focus group, 20 students voted on the DPV framework's clarity: 6 found it "easy to understand," while 14 voted "just right," with none indicating "difficult." These results suggest that the framework is appropriately designed for middle school students, enabling them to grasp and interpret its concepts effectively. Students acknowledged that while the terminology was initially challenging to remember, they found it easier as they engaged more deeply with the framework. As one noted, "After we talked more about it and used it, I could just remember it all" (FG3-2). Others found it intuitive, as one explained, "It felt like you'd already learned some of this before in school" (FG1-1, FG1-2). This familiarity may stem from the use of vocabulary and a worksheet style that mirrored their school experiences—such as terms like "Problem statements" and "Big Questions." These familiar elements likely helped activate their prior knowledge and facilitated understanding, even though none of the students had previous experience with data visualization.

Students were also asked to explain the DPV framework as if the interviewers were unfamiliar with it. In each group, one student voluntarily started, and others contributed. All groups correctly identified D for Domain, P for Purpose, and V for Visual, though two groups referred to "Visual" as "Visualization." They accurately recalled sub-questions for each stage and their roles in the design process. For example, Domain was described as "what people want to know about the information" (FG4-2), Purpose as "to show the story of comparison/distribution/composition/relationship" (FG2-5), and Visual as "how to make different variables into parts of the chart" (FG1-1). These responses show students could accurately interpret and articulate the DPV framework, meeting **DG2**.

What do students think of the positive aspects of the DPV? To understand students' opinions on using the DPV framework in creating data visualization, we analyzed their responses to the question, "Do you think the DPV framework is helpful for creating data visualizations, and why?" using an inductive thematic analysis approach [BC06b]. Through the analysis, we identified five themes related to the value of the DPV framework.

Many students praised the DPV framework for simplifying the complicated data visualization design process, meeting **DG2&4**. For example, FG1-1 commented that "DPV is like a set of materials, it guided you through your visualization process". FG3-3 noted, "Because it helped me break down each part and (I) doesn't get overwhelmed," illustrating how dividing the design into smaller sections aids understanding and retention. Six students viewed the DPV framework as a set of guiding questions to promote their thinking. FG4-1 mentioned, "when I don't know what to do, and when I use DPV, ..., I can come up with a few questions, and those helped me." FG2-3 also commented during their project-creating process that, "I do think that it will try to map out and think that we wanted to do before we actually started", indicating that the DPV framework helps students brainstorm and strategically plan projects, an essential data science skill for real-world challenges. Two students (FG4-1, FG4-4) noted that besides data visualization challenges, the DPV framework could aid them "in math ... to see what type is right for this purpose of whatever project I'm doing", and "when we want to know about our project, we could just think of DPV, and then we'll have a start already", indicating students' intention to transfer the DPV framework to other domain. Students found the DPV framework useful for comparing design ideas and enhancing creativity, meeting **DG3**. FG3-4 said, "It helps me be more creative and also just recognize how much like what I should be doing." It also encouraged considering the audience's perspective, as FG3-1 stated, "It helps you break down (vis) so the audience can, like actually can find it helpful in many ways." Furthermore, it also expanded students' knowledge to explore things that are unknown to them, as FG3-2 remarked "It helped me know about DPV a little bit more than I do now than I do than I did before."

What do students think is a negative aspect of the DPV? We also ask them, "Which part of the framework does not make sense to you?" Three students felt the DPV pedagogy had too much information, claiming that "there are too many terminologies." They also commented on the overall DPV pedagogy that "there is a lot of paperwork," expressing a preference for online tools like Google Docs for easier organization of learning materials. Students also provided feedback on each element of the DPV framework:

- **Domain:** Two students regarded Domain as hard to understand. Both of them pointed out that it was still hard for them to distinguish between "convention" and "visual." They both acknowledged that they "can tell the difference but cannot find a better way to self-explain (the difference)."
- **Purpose:** Seven students pointed out that they have some difficulties in understanding concepts related to the four types of Big Idea when articulating their purpose. Specifically, two students pointed out that it was hard to "figure out which chart goes with which type (of the Big Idea)" because "one chart might apply to multiple." Regarding the four types of Big Idea, four students regarded the "Relationship" type Big Idea as hard to understand, with one student suggesting the use of "connection" instead of "relationship" to describe this scenario. They also regarded "Composition" as hard to understand.
- **Visual:** While no students regarded Visual as hard to understand, one student highlighted that since we listed several channels (color, size, position, etc.) in the worksheet, they felt constrained by the item list because "shape and others can also be visual."

6. Discussion

6.1. Age-Appropriate Adaptation

Adapting existing pedagogical frameworks to a K-12 context required simplifying its terminology usage to improve students' comprehension and engagement. In this work, efforts were spent to adapt the existing Four-level Bested Model [Mun09]to K-12 scenarios and found to be effective. For example, renaming "Abstraction Design" to Purpose and "Encoding" to Visual made the framework more accessible by aligning with students' everyday vocabulary at school, reducing cognitive load, and focusing attention on core concepts. Introducing familiar terms like "Big Idea" further bridged abstract concepts with practical applications, enhancing understanding and also activating students' prior knowledge. However, "Convention" remained challenging for students to grasp. Attempts to explain it as "rules" or "best practices" were insufficient. A local science teacher suggested using the analogy of traffic lights, where red and green have defined meanings, versus holiday lights, where colors are chosen for aesthetics. This example highlights the importance of grounding explanations in relatable contexts and scenarios that students are more familiar with to make complex ideas clearer, which is also discussed in [JRSD*21] of leveraging learners' personal experience to promote data literacy education. These adaptations emphasize the need for continuous refinement of language to enhance accessibility and learning outcomes.

6.2. The DPV Framework for critique

The DPV framework not only guides students in designing data visualizations but also helps them critique existing ones, a key design goal as addressed in **DG3**. This aligns with the stages of data visualization in the revised Bloom's Taxonomy [Fry07], covering both lower- and higher-order thinking levels.

At lower-order levels, students analyze visualizations by evaluating their alignment with domain conventions, clarity of purpose, and effectiveness of visual encoding. For instance, they assess whether a visualization accurately represents data relationships, adheres to domain-specific conventions, or effectively uses color, size, and position. At higher-order levels, the DPV framework encourages students to apply, evaluate, and redesign visualizations [VW15]. By reflecting on design choices and identifying areas for improvement, students develop critical thinking and analytical skills. The DPV framework's structured approach ensures students systematically engage with visualizations at all three levels, enhancing their ability to critique existing visualizations while reinforcing design principles they can apply to their own work.

6.3. Empowering K-12 Learners with Data Visualization

Engaging K-12 students and teachers is crucial for empowering the next generation in today's data-driven world [BZP*19, DB18]. This work also aims to broaden participation in the data visualization community by designing the DPV framework with direct input from educators and students, ensuring it addresses their authentic needs and unique challenges. Teachers contributed valuable suggestions to align the framework with curriculum standards and student learning needs, guiding the development of practical resources like worksheets and activities. Student feedback was equally important, leading to improvements in terminology and step-by-step guidance for accessibility. By incorporating these perspectives, the DPV framework bridges the gap between professional data visualization practices and K-12 education. It not only equips teachers and students with data skills but also fosters a sense of belonging in the broader data visualization community. This inclusive approach enriches the community with fresh perspectives, cultivates diverse and inclusive participation [BKR*23, LRCT22, SOX*24], and ultimately promotes equity and innovation in the field.

6.4. Limitation and future iteration

Student feedback highlighted several limitations in the DPV framework, identifying areas for improvement. At the Domain level, students struggled to distinguish between "convention" and "visual," suggesting a need for clearer explanations and more concrete examples. At the Purpose level, students struggled with the four types of Big Ideas and matching chart types to specific categories due to their overlap. This highlights the inherently iterative nature of data visualization design, where rigid categorization can limit flexibility. While the decision tree offers valuable guidance for selecting appropriate chart types, it has limitations when applied to real-world data, which often involves multiple purposes. Recognizing these challenges, we encourage students to use the decision tree as a starting point while remaining adaptable in their design process to better align with their goals and data. At the Visual level, one student felt constrained by the explicit list of visual channels, highlighting a trade-off between scaffolding the design process and allowing open-ended interpretation. This feedback underscores how the DPV framework engages students in a core aspect of visualization: understanding how visual elements map to data properties. This aligns with the framework's goal of fostering authentic design experiences and improving data visualization skills. Future iterations of the DPV framework should incorporate iterative design processes, offer more nuanced explanations, and provide opportunities for students to adapt and expand upon the framework as their understanding grows. These refinements aim to make the framework both more effective and adaptable to diverse learning needs.

7. Conclusion

This paper presents the DPV (Domain, Purpose, Visual) framework, designed to scaffold K-12 students through the data visualization design process. The framework was derived based on existing design frameworks for higher education and K-12 teachers with experience teaching data visualization. Evaluated during a summer camp with middle schoolers, the framework proved effective in helping students create visualizations to support communication. Feedback from the evaluation informed iterative refinements, ensuring the framework remains effective in bridging professional visualization practices with K-12 contexts. By empowering K-12 teachers and students, the DPV framework contributes to broadening participation in data visualization and fostering data literacy.

8. Acknowledgments

This material is based upon work supported by the National Science Foundation (NSF AISL #2314109).

References

[Abe21] ABELA A.: Step 7 graphics: Identify the most effective graphical elements to use in your presentation, Jan 2021. URL: https://extremepresentation.com/design/7-charts/. 5

[BBG19] BÖRNER K., BUECKLE A., GINDA M.: Data visualization literacy: Definitions, conceptual frameworks, exercises, and assessments. *Proceedings of the National Academy of Sciences 116*, 6 (2019), 1857–1864. doi:10.1073/pnas.1807180116. 1

[BC06a] BRAUN V., CLARKE V.: Using thematic analysis in psychology. *Qualitative research in psychology 3*, 2 (2006), 77–101. 4

[BC06b] BRAUN V., CLARKE V.: Using thematic analysis in psychology. *Qualitative Research in Psychology 3*, 2 (2006), 77–101. doi:10.1191/1478088706qp063oa. 7

[BKR*23] BACH B., KECK M., RAJABIYAZDI F., LOSEV T., MEIRELLES I., DYKES J., LARAMEE R. S., ALKADI M., STOIBER C., HURON S., ET AL.: Challenges and opportunities in data visualization education: A call to action. *IEEE Transactions on visualization and computer graphics* (2023). doi:10.1109/TVCG.2023.3327378. 1, 8

[BMBH16] BÖRNER K., MALTESE A., BALLIET R. N., HEIMLICH J.: Investigating aspects of data visualization literacy using 20 information visualizations and 273 science museum visitors. *Information Visualization 15*, 3 (2016), 198–213. doi:10.1177/1473871615594652. 1, 2

[BSK*24] BOUCHER M., STOIBER C., KECK M., DE-JESUS-OLIVEIRA V.-A., AIGNER W.: The comic construction kit: An activity for students to learn and explain data visualizations. In *2024 IEEE Visualization and Visual Analytics (VIS)* (2024), IEEE, pp. 86–90. doi:10.1109/VIS55277.2024.00025. 1, 2

[BYP21] BEYER J., YANG Y., PFISTER H.: Visualization design sprints for online and on-campus courses. *IEEE Computer Graphics and Applications 41*, 6 (2021), 37–47. doi:10.1109/MCG.2021.3115413. 1, 2

[BZP*19] BISHOP F., ZAGERMANN J., PFEIL U., SANDERSON G., REITERER H., HINRICHS U.: Construct-a-vis: Exploring the free-form visualization processes of children. *IEEE transactions on visualization and computer graphics 26*, 1 (2019), 451–460. doi:10.1109/TVCG.2019.2934804. 8

[DB18] D'IGNAZIO C., BHARGAVA R.: Creative data literacy: A constructionist approach to teaching information visualization. 8

[Dea14] DEAHL E.: Better the data you know: Developing youth data literacy in schools and informal learning environments. *Available at SSRN 2445621* (2014). doi:10.2139/ssrn.2445621. 2

[FBL24] FERNANDEZ C., BLIKSTEIN P., LOPES R. D. D. L.: Design failures in data visualization programming activities. In *Proceedings of the 23rd Annual ACM Interaction Design and Children Conference* (2024), pp. 574–586. doi:10.1145/3628516.3663348. 2

[Fry07] FRY B.: *Visualizing data: Exploring and explaining data with the processing environment*. " O'Reilly Media, Inc.", 2007. 2, 8

[HA17] HE S., ADAR E.: Vizitcards: A card-based toolkit for infovis design education. *IEEE Transactions on Visualization and Computer Graphics 23*, 1 (2017), 561–570. doi:10.1109/TVCG.2016.2599338. 1, 2

[HBH*20] HURON S., BACH B., HINRICHS U., KECK M., ROBERTS J. C.: Ieee vis workshop on data vis activities to facilitate learning, reflecting, discussing, and designing. In *IEEE VIS 2020* (2020). 1

[IFMTW23] ISRAEL-FISHELSON R., MOON P. F., TABAK R., WEINTROP D.: Preparing students to meet their data: an evaluation of k-12 data science tools. *Behaviour & Information Technology* (2023), 1–20. doi:10.1080/0144929X.2023.2295956. 1

[JRSD*21] JOHNSON B., RYDAL SHAPIRO B., DISALVO B., ROTHSCHILD A., DISALVO C.: Exploring approaches to data literacy through a critical race theory perspective. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (2021), pp. 1–15. doi:10.1145/3411764.3445141. 8

[KHP*23] KECK M., HURON S., PANAGIOTIDOU G., STOIBER C., RAJABIYAZDI F., PERIN C., ROBERTS J. C., BACH B.: Eduvis: Workshop on visualization education, literacy, and activities. *arXiv preprint arXiv:2303.10708* (2023). doi:10.48550/arXiv.2303.10708. 1

[LDCT21] LEE V. R., DRAKE J., CAIN R., THAYNE J.: Remembering what produced the data: Individual and social reconstruction in the context of a quantified self elementary data and statistics unit. *Cognition and Instruction 39*, 4 (2021), 367–408. doi:10.1080/07370008.2021.1936529. 2

[LRCT22] LOSEV T., RAYNOR J., CARPENDALE S., TORY M.: Embracing disciplinary diversity in visualization. *IEEE Computer Graphics and Applications 42*, 6 (2022), 64–71. doi:10.1109/MCG.2022.3209605. 8

[MLM17] MCKENNA S., LEX A., MEYER M.: Worksheets for guiding novices through the visualization design process. *arXiv preprint arXiv:1709.05723* (2017). doi:10.48550/arXiv.1709.05723. 1, 2

[MMAM14] MCKENNA S., MAZUR D., AGUTTER J., MEYER M.: Design activity framework for visualization design. *IEEE Transactions on Visualization and Computer Graphics 20*, 12 (2014), 2191–2200. doi:10.1109/TVCG.2014.2346331. 2

[MSQM15] MEYER M., SEDLMAIR M., QUINAN P. S., MUNZNER T.: The nested blocks and guidelines model. *Information Visualization 14*, 3 (2015), 234–249. doi:10.1177/1473871613510429. 2

[Mun09] MUNZNER T.: A nested model for visualization design and validation. *IEEE transactions on visualization and computer graphics 15*, 6 (2009), 921–928. doi:10.1109/TVCG.2009.111. 2, 4, 8

[POPR16] PHILIP T. M., OLIVARES-PASILLAS M. C., ROCHA J.: Becoming racially literate about data and data-literate about race: Data visualizations in the classroom as a site of racial-ideological micro-contestations. *Cognition and Instruction 34*, 4 (2016), 361–388. doi:10.1080/07370008.2016.1210418. 2

[RHR15] ROBERTS J. C., HEADLEAND C., RITSOS P. D.: Sketching designs using the five design-sheet methodology. *IEEE transactions on visualization and computer graphics 22*, 1 (2015), 419–428. doi:10.1109/TVCG.2015.2467271. 2

[RRJH17] ROBERTS J. C., RITSOS P. D., JACKSON J. R., HEADLEAND C.: The explanatory visualization framework: An active learning framework for teaching creative computing using explanatory visualizations. *IEEE transactions on Visualization and Computer Graphics 24*, 1 (2017), 791–801. doi:10.1109/TVCG.2017.2745878. 2

[Shr18] SHREINER T. L.: Data literacy for social studies: Examining the role of data visualizations in k–12 textbooks. *Theory & Research in Social Education 46*, 2 (2018), 194–231. doi:10.1080/00933104.2017.1400483. 1

[SMM12] SEDLMAIR M., MEYER M., MUNZNER T.: Design study methodology: Reflections from the trenches and the stacks. *IEEE transactions on visualization and computer graphics 18*, 12 (2012), 2431–2440. doi:10.1109/TVCG.2012.213. 2, 4

[SOX*24] SEO J. Y., O'MODHRAIN S., XIA Y., KAMATH S., LEE B., COUGHLAN J. M.: Designing born-accessible courses in data science and visualization: Challenges and opportunities of a remote curriculum taught by blind instructors to blind students. doi:10.2312/eved.20241053. 8

[VW15] VIÉGAS F., WATTENBERG M.: Design and redesign, Apr 2015. URL: https://medium.com/@hint_fm/design-and-redesign-4ab77206cf9. 8

[WGM*16] WOLFF A., GOOCH D., MONTANER J. J. C., RASHID U., KORTUEM G.: Creating an understanding of data literacy for a data-driven society. *The Journal of Community Informatics 12*, 3 (2016). doi:10.15353/joci.v12i3.3275. 2