DocuBits: VR Document Decomposition for Procedural Task Completion

Geonsun Lee*
University of Maryland, College Park

Jennifer Healey†
Adobe Research
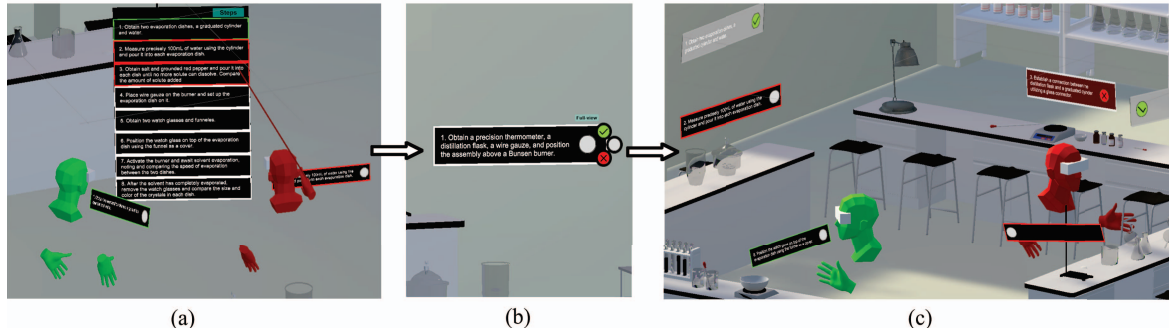
Dinesh Manocha‡
University of Maryland, College Park

Figure 1: DocuBits: Interactive document fragments strategically anchored to pertinent task locations, minimizing attention shifts between reading and task execution, and fostering collaborative learning. In a multi-user context, participants can (a) fragmentize and select specific sections of text instructions for execution, (b) interact with the DocuBit to monitor progress, and (c) position these fragments within the virtual environment for seamless sharing among all users.

ABSTRACT

Reading monolithic instructional documents in VR is often challenging, especially when tasks are collaborative. Here we present DocuBits, a novel method for transforming monolithic documents into small, interactive instructional elements. Our approach allows users to:(i) create instructional elements (ii) position them within VR and (iii) use them to monitor and share progress in a multi-user VR learning environment. We describe our design methodology as well as two user studies evaluating how both individual users and pairs of users interact with DocuBits compared to monolithic documents while performing a chemistry lab task. Our analysis shows that, for both studies, DocuBits had substantially higher usability, while decreasing perceived workload ($p < 0.001$). Our collaborative study showed that participants perceived higher social presence, collaborator awareness as well as immersion and presence ($p < 0.001$). We discuss our insights for using text-based instructions to support enhanced collaboration in VR environments.

Index Terms: Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Virtual reality; Human-centered computing—Human computer interaction (HCI—Interaction techniques; Human-centered computing—Human computer interaction (HCI—Interaction paradigms—Collaborative interaction;

1 INTRODUCTION

Virtual Reality (VR) can provide a safe place for people to learn how to perform tasks that would be difficult, dangerous, or expensive to practice in the real world, e.g., flight simulation, surgery, cooking, or laboratory experiments. The most common way that people learn to perform a task is with the help of a set of written instructions. Even when augmented by live teacher demonstration, written instructions provide a common standard for how a process should be completed.

*e-mail: gsunlee@umd.edu
†e-mail: jehealey@adobe.com
‡e-mail: dmanocha@umd.edu

Many institutions seek a standardized way of training workers and many school systems have a common core of learning to ensure similar instruction across schools. Our goal is to provide a method to retain all the benefits of this kind of standardized instruction while completing it and extending instruction to virtual environments.

Currently, most virtual reality instruction is customized to the application with specialized follow-along demonstrations incorporated into the virtual environment [36]. This kind of instruction is generally preferred because it is simple to follow and takes advantage of the richness of the virtual world.

We note, however, that there are two possible drawbacks to this approach. The first is that some processes already have set of text instructions used in physical lab training and that these might be periodically updated. With DocuBits, the current, approved text can be directly brought into the VR environment without having to re-make a demonstration video. The second is that demonstration might inspire mimicry rather than challenging the learner to reason about the instruction in an abstract way. We believe that the benefit of text-based instructions is that they are identical across physical and virtual environments, they are abstract, and they remain human-readable indefinitely.

Reading text in VR also presents its own set of challenges, especially when text is longer and when tasks are shared. These include: reading itself, task switching, tracking progress and communicating progress. [18, 29, 30]. DocuBits addresses the challenge of reading longer text by simply breaking up the text into smaller units. These smaller text units are then placed withing the interactive DocuBit and can be placed strategically within the environment. With smaller bits of text, the reader can read once and simply glance back to refresh their memory about the current task step, allowing them to transition more freely between reading instructions and doing the work. The interactive aspect of DocuBits helps readers keep track of their work with different display modes for unattempted instructions, completed tasks and failed or blocked tasks. These different displays also simplify communication between users working on shared tasks.

We believe that this interactive display aspect is key to the value of DocuBits and that it will allow multiple participants to share in both reading and doing tasks equally as well as provide an overall "birds-eye" view of progress to a third party, such as a teacher, who might want to oversee the users experience. This is an improvement

to some current lab situations where the problem of task switching is solved by one person taking a "reader" role and another taking a "doer" role (resulting in an uneven experience) and where teachers must go around to each group to assess progress instead of being able to get a quick overall view.

Main Results: We address the challenge of following procedural documents in VR with DocuBits, a novel method for transforming monolithic instructional documents into text fragments and embedding those text fragments into portable interactive virtual objects. These virtual text elements (text segments) can be moved and placed next to virtual task spaces that comprise both active elements (multi-state checkboxes), which specifically convey the state of the task, and state-based behaviors (floating up, fading away, bouncing up and down), which communicate the current state of a task from a distance (especially beneficial for multi-user collaboration). Not only does the user have a sense of action history, but collaborators can proactively support each other as the progress is visually shared.

In this paper, we present our methodology for the design and development of DocuBits, a technical description of our system, and two user evaluations, one with an individual user and one with pairs of users who interact both with the DocuBits and each other. We use a fictional chemistry lab experiment as a testbed for our evaluations while envisioning the broad applicability of our method to any procedural task suitable for VR instruction. Our evaluations show that DocuBits both improves the immersion and lowers the cognitive load when compared to a baseline of non-segmented, monolithic documents of instructions. In the paired-users evaluation, participants report higher perceived levels of co-presence and collaboration. This demonstrates DocuBits' potential to improve the effectiveness of VR training for procedural tasks and enable multi-user collaboration. In summary, our contributions include:

- Uncovering insights from real-life instructors on the challenges associated with text instructions in educational settings, thereby informing design considerations for VR interfaces.

- Presenting DocuBits, an innovative document interface that allows fragmentation, interaction, and anchoring within the VR space, enhancing the effectiveness of reading-while-doing.

- Introducing four user interactions: (i) Doc to Bits; (ii) Tag Along and Stick; (iii) Progress animation; and (iv) Assignment, which are combined for the overall interactive approach.

- Conducting a comprehensive user study on DocuBits, encompassing both single-user and paired-user scenarios.

- Demonstrating the significant impact of DocuBits on user experience, particularly in collaborative learning scenarios, and extracting valuable insights on user behavior with text instructions from the study.

2 Related Works

2.1 VR for Training and Education

Virtual reality has often been used to train people to perform tasks in cases where tasks are dangerous, difficult, or expensive to replicate. Examples of this include high-impact activities such as surgical work, firefighting [43], first responder training [41], aircraft and spacecraft control [34], and construction work [1]. Additionally, VR has proven effective in more routine yet skill-intensive tasks such as assembly and maintenance. For instance, VR training systems have been developed for automotive assembly tasks [4, 20], offering a safe and cost-effective method for training workers in complex assembly processes. In the field of maintenance, VR has been used to train technicians in the maintenance of industrial equipment, reducing the risks and costs associated with hands-on training [2, 21]. This flexibility of virtual reality also extends to other common tasks such as cooking [19], Tai Chi [8], and running a chemistry laboratory [22], where time and materials make repeated practice costly.

VR's on-demand instruction allows users to practice tasks flexibly, free from traditional classroom constraints. This adaptability supports individual learning needs and pacing [36]. Its key advantage lies in reducing social pressures for better skill focus and lowering resource use compared to physical training environments. This is particularly beneficial in scenarios with limited face-to-face interaction [46].

For educational applications of VR, Dieterle and Clark created River City, a virtual environment for fostering scientific inquiry and skills in middle school science classes [12]. Stokes et al. investigated attention distraction in children with ADHD within a simulated VR classroom, employing eye-tracking measures to overcome traditional laboratory method limitations [42]. Parsons et al. utilized a VR classroom setting, investigating its potential for controlled performance assessments in an ecologically valid context, highlighting the impact of distraction stimuli when contrasted with standard neuropsychological methods [35].

2.2 Collaborative Work in VR/AR

Prior work has proven the positive impact of VR and AR in collaborative and distance education settings [7]. Simeone et al. [40] compared a system-based teaching system with a single-user and a two-user version; the latter had one user acting as a teacher. The study results revealed that the two-user version scored higher in overall preference and clarity compared to the single-user version, which used animation sequences for teaching.

Radu et al. [37] found that in a robot peer-programming task, augmenting informatic visualizations through AR headsets led to higher team learning gains compared to learning without visualizations. Some works researched collaborative learning scenarios where users use different devices and have different roles assigned. For example, a table is given to a user who has the guiding role within an educational game, similar to that of a teacher [32, 44, 47]. Drey et al. [17] developed VR pair-learning systems and demonstrated that it is preferable to have a symmetric system, with both participants wearing a VR HMD, than an asymmetric system, with one student using a tablet and the other a VR HMD, in terms of immersion, presence, and cognitive load. However, these works neglect the role of text presentation in their scenarios. Additionally, they mostly consider a teacher-student relationship rather than a peer-to-peer learning scenario which is what DocuBits focuses on supporting.

Schott et al. [39] built a VR/AR multi-user environment for liver anatomy education, considering how certain types of information should be presented in the environment, such as visualizing 3D models in a shelf-like object and presenting text and image data on a shared board. Jin and Lie et al. [28] investigated collaborative learning on VR video viewing systems, comparing viewing modes based on whether users have individual control or shared control over video sync. Results showed that shared VR video modes increase collaboration and social presence. Loki [45] is a system that supports remote instructions by capturing video, audio, and spatial information for a mixed-reality presentation of the learner and the listener. It allows learners to record their performance so that a remote instructor can review and annotate them in AR and VR. While these works have concentrated on demonstration-led learning, DocuBits shifts the focus to the often-overlooked importance of text-based instruction, enhancing learning where following the documentation is key.

2.3 Text Presentations and Interfaces

Investigating legibility in text presentation on electronic devices has been a longstanding concern, with Dillon et al. [14] addressing its importance. In virtual environments, text presentation encounters unique challenges due to resolution constraints and the introduction of a third dimension. Chen et al. [6] evaluated different combinations of VR navigation techniques and text layout techniques in

Information-Rich Virtual Environments. Dittrich et al. [16] proposed guidelines for text visualizations in 3D virtual spaces, emphasizing the need for larger text sizes than those on 2D displays. Jankowski et al. [26] explored the integration of text with video and 3D graphics, revealing preferences for negative text presentations. Dingler et al. [15] investigated optimal parameters for VR text presentation, including text size, convergence, and color. Recent efforts include studies on text presentations on 3D objects with various surfaces [48], revealing the ease of reading text wrapped around a 3D object with a single axis.

While reading performance on electronic devices has been extensively compared [5, 9, 11, 13, 23], Rau et al. [38] evaluated reading speed and accuracy in VR and AR, finding differences compared to traditional computer screens. Some works investigated optimizing interfaces for reading long texts in VR. Kojic et al. [30] tested different values for text parameters such as font size, distance, type of HMDs used, etc., for short, medium, and long text samples. Kobayashi et al. [29] explored text parameters for reading long texts and proposed view settings for better readability and less fatigue in VR. Gabel et al. [18] compared four text panel UIs to present long texts, finding no significant differences in reading performance. Lee et al. [31] developed a set of gaze-based interactions to improve the user's reading experience in VR, showing enhanced effectiveness and less perceived workload.

However, these works only consider the sole activity of reading and not how users interact with texts when they are performing tasks based on them. DocuBits proposes how we can support the user in such experiences and even multi-user scenarios for this.

3 INITIAL CONCEPT AND DESIGN

Our concept focuses on fragmenting documents into smaller units, strategically placed in VR to enhance text accessibility. To refine this concept, we consulted experienced educators for insights into learners' needs in instructional settings.

3.1 Needs-Finding Study

To develop a set of design principles for DocuBits beyond the idea of segmented text we recruited a focus group of 7 experienced educators in three different areas. The areas of expertise include physics (E1), biochemistry (E2, E3), computer science (E4, E5), jewelry craft (E6), and landscape design (E7). All the participants had 3 to 6 years of experience in teaching students and trainees and often followed instructional documents such as class projects and laboratories. All participants spoke the same language, and the focus group interview was conducted in two separate sessions to accommodate their availability.

During the interviews, we focused on several key areas: methods of distributing instructional documents to students, observations of student interactions and collaboration with these documents, frequency of students referring to instructions during tasks, and educators' involvement in student help requests. We also inquired about students' collaborative dynamics, such as role division in group tasks, and identified the main challenges students encountered while following procedural instructions, including understanding the written content and managing collaborative efforts.

Key takeaways included the educators' use of both physical and digital documents and the dynamics of student interactions when using reading materials in group tasks. Common queries from students were about locating necessary objects and the practical difficulties in handling documents in lab settings, such as exposure to fluids and managing them with gloves or protective eyewear.

We present the results of the interview study in terms of the benefits and drawbacks of utilizing documents in instructing students in the following paragraphs.

3.2 Benefits of Utilizing Text Instructions

Self-paced learning All educators agreed that they distributed instructional documents instead of demonstrating each step in a process because a document provides a concrete reference for the steps. Demonstrations often demand continuous viewing up to a specific time step for comprehension and necessitate re-watching from the beginning if details are forgotten, making them more time-consuming compared to text-based instructions. E2, E3, and E6 pointed out that they usually do a demonstration of the process to the class and then allow students to go at their own pace afterward by referencing and following the steps in the document.

Content augmentation and reformatting To mitigate situations where the students lose track of their own progress or need their own interpretation to help them perform the task, students augment the document with actions such as adding drawings and crossing out steps that are finished or highlighting parts that need to be revisited. When accessing digital documents, there are cases in which students go further by editing and creating their own documents so the task is decomposed or integrated in a way that is more convenient for their own workflow.

Overall, although the interviewers unanimously recognized the advantages of employing distributed instructional documents for teaching students, there is a notable dearth of tools available to facilitate the efficient utilization of such documents by students. This issue becomes even more pronounced in collaborative learning scenarios.

3.3 Drawbacks in Utilizing Text Instructions

Losing track of progress The majority of the educators (E2, E3, E4, E5) noted that students frequently get "lost" in their progress when following documents, particularly when switching between tasks, changing positions, or asking questions. E5 remarked, "When a student had a question to ask, there was a time gap until I came over, and then the student had to skim through the document again to find the point they wanted to refer to." E2 added, "This happens even when they are following the document themselves, as I once observed a student redoing steps to trace back their mistakes."

Unequal role distribution of reading and doing In collaborative experiments, E2 and E3 observed scenarios where one student read out instructions for another performing the task. E1, E4, E5, and E6 noted in larger group projects, often one student would take charge of tracking progress, leading to an unequal distribution of understanding among the team. This was a concern for five of the interviewees as it could lead to a skewed comprehension of the task across the group.

Credit Attributing Instructors who incorporated physical movement within instructional procedures (E2, E3, E6, E7) highlighted difficulties in identifying which student was responsible for specific tasks, especially when using online documents. This often made it challenging for instructors to offer timely assistance and for peers to support each other effectively in the learning process.

In conclusion, while instructional documents offer advantages such as personalized learning and progress tracking, there are notable challenges in managing student understanding, collaboration, and prompt assistance in collaborative learning scenarios. We use these observations to design our interface in Section 4.

3.4 Design Considerations for Docubits

Based on the group interview insights, we identified key design principles for DocuBits, focusing on the need for collaborative tools in instructional document use.

Dynamically fragment documents: Considering diverse learning paces, it is suggested to design DocuBits to allow students to dynamically fragment documents into smaller segments like numbered steps or sentences. This flexibility is vital for catering to both
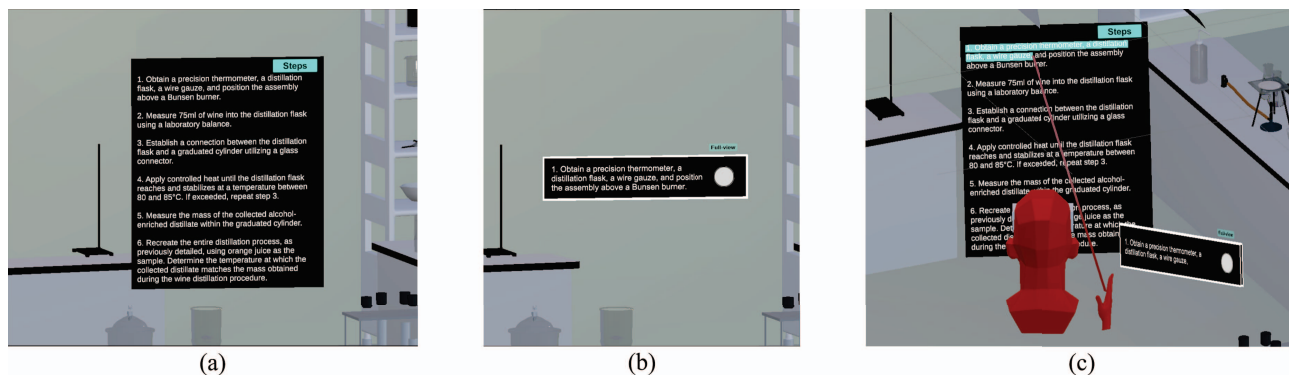
Figure 2: Our method for creating Doc to Bits. (a) The initial monolithic document appears, and users can either (b) create DocuBits with pre-defined numbered steps or (c) customize DocuBits by highlighting and releasing words.

individual preferences and group collaboration needs, allowing for customized information flow.

Spatial control: Given the varying spatial dynamics in learning environments, DocuBits should enable users to control the positioning of documents effectively. This feature would ensure that whether students are stationary or moving, they have continuous and easy access to instructional content.

Visibly share progress: Incorporating interactive elements within documents, such as checklists and drawings, is recommended for efficient progress tracking. The implementation should include visible indicators of task status and progress, which would enhance learning experiences by minimizing the need for verbal communication in both individual and group settings.

4   DOCUBITS: INTERACTIVE SPATIALIZED TEXT INSTRUCTIONS

In accordance with our design considerations, we developed and implemented DocuBits, an interactive solution for users to effectively engage with instructional documents in VR training scenarios. Our goal was to create a versatile framework that worked well for both individual and group-based learning contexts. We wanted to implement and test several experiences that we believed would address the needs uncovered in our formative study. To this end, we enabled four key experiences: (i) Doc to Bits a method for fragmenting a procedural document into cohesive units of texts representing tasks; (ii) Tag Along and Stick; a method for easily placing those bits in the environment; (iii) Progress Animation, a method for visibly expressing task progress; and (iv) Assignment, a method for associating bits to a particular user.

Our approach has been developed using the Unity game engine with Photon [10] for multi-player networking. To deliver an immersive VR experience, we use two Oculus Quest 2 head-mounted displays (HMDs) per user. The virtual environment has been meticulously set up to resemble a chemistry lab, featuring scattered equipment within the spatial domain.

4.1   Doc to Bits

We enable two experiences for users to convert monolithic documents into bits, one that automatically segments the document by numbered steps and one that allows custom separation by highlighting. Upon entering the VR environment, the instructional documents initially appear as a regular "monolithic" document (Figure. 2-(a)). In our use cases, all documents had numbered steps, but we believe many procedural documents with other kinds of separators, dashes, dots, lines, etc. can be converted into documents with numbered steps in a straightforward manner. Using the first approach, the user simply clicks on a button labeled "steps" and the monolithic document automatically transforms into a stack of DocuBits, one

for each numbered instruction (Figure. 2-(b)). The second approach is custom highlighting. This approach is useful when users want more granular control over what parts of the document should be individual DocuBits (Figure. 2-(c)). For example, some documents may contain long steps that might be better as two or three DocuBits, while some steps might include more text than is necessary for a DocuBit. Highlighting offers this fine-grain editing control while still only requiring the user to swipe over the desired text using a press-and-drag action. Upon releasing the controller button, the highlighted segment is recognized as a single DocuBit. All created DocuBits are displayed as stacks that remain fixed to the user's avatar position, facilitating ease of movement within the VR environment. Note that users can switch to a "full view" of the document as they desire. Figure 2 shows the method and the user's views.

4.2   Tag-Along-and-Stick

Once a stack of DocuBits is created, users can sort the stack and then take any individual Docubit, stack, or sub-stack with them. The DocuBits "tag along" behind the person's avatar as they navigate the space. The user is then free to grab a DocuBit and place it in the environment where they feel the instructions would be most helpful. Once placed, the DocuBit positions can be saved with the environment. This would enable, for example, a teacher to set up the DocuBits for students in advance of a lesson to save time. It also enables students to save progress if procedures need to be interrupted and resumed later.

In DocuBits, users can switch between different DocuBits and affix them within the virtual environment. When a DocuBit is placed or "stuck" in a specific location, it remains stationary. However, if the corresponding task associated with the DocuBit is still in progress (not marked as finished), an additional DocuBit clone is generated and follows the user when the initially placed DocuBit is no longer within the user's view frustum (Figure. 3-(c)). Furthermore, users can revisit previously encountered DocuBits by flipping through them. These past Docubits are visually distinguished by a gray tint, indicating their completed status.

4.3   Progress Animation

Each DocuBit also has a status indicator and associated animation behaviors that show if a task is not attempted, attempted but not completed, or successfully completed (Figure. 3-(a)). Steps that have not yet been attempted have a clearly visible white color and are placed near a task. Upon successful completion, a green light appears in the status indicator circle, the body of the object turns gray, and the object floats up (Figure. 3-(d)). If the person cannot complete the task (for example, the user has a question about the instructions or cannot determine how the instructions relate to the virtual world) the user can toggle the status indicator to red and the DocuBit will present with a red indicator and fully visible text (Figure. 3-(b)).
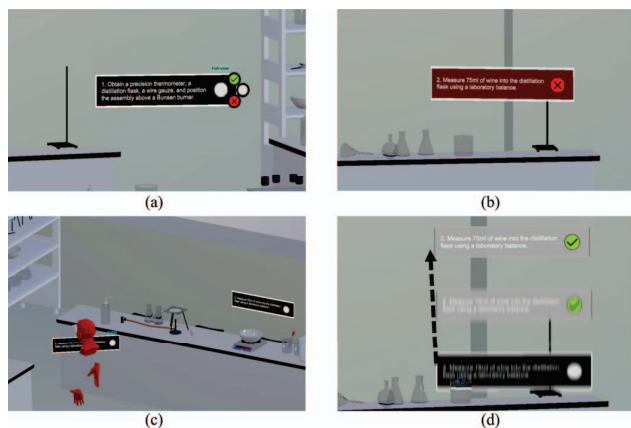
Figure 3: DocuBits are interactive text elements that maintain and display the status: (a) Users can select a status such as: (b) task attempted but not successfully completed, (c) task in progress, or (d) task successfully completed. When a DocuBit is still in progress, a clone of a DocuBit will tag along with the user as in (c). Depicted in (d) is the animation of a complete DocuBit floating upwards.

These interactive behaviors are designed to capture users' attention intuitively regarding users' overall progress. While the completion of an action can be automatically detected depending on the task scenario, in our implementation we allowed users to manually select the status of the DocuBit. Such indications are valuable for both multi-participant task teams and for teachers who are overseeing the progress of multiple students in the virtual classroom.

4.4 Assignment

In collaborative settings, when multiple users are performing a task together, the assignment of DocuBits to individual users becomes crucial. To denote ownership, distinct color codes are employed, such as red for User A and green for User B. The ownership allocation occurs during the initial fragmentation of the monolithic document. If users opt to fragment the document automatically based on pre-defined steps (Figure. 4-(a)), one participant would need to specify how many people will be participating. Then the selected blocks of content will be divided into task-cohesive stacks in front of each user. Currently, task-cohesive steps need to be specifically identified by the document author, but ideally, these could be logically inferred in some future implementation. If users choose to create DocuBits by highlighting specific words (Figure. 4-(b)), each DocuBit is assigned to the participant who created it after the user selects and releases the trigger button to highlight a sentence.

If users wish to change the ownership of a DocuBit while performing a task, they can easily modify ownership by selecting the frame of the DocuBit. Once the ownership is altered, the DocuBit seamlessly transitions into the new owner's stack of DocuBits (see Figure. 5). Note that DocuBits that are marked "completed" cannot be applied for re-assigning ownership.

5 EVALUATION STUDY

In this section, we present the experimental design for evaluating the effectiveness of DocuBits in two distinct scenarios. First, we assess the standalone performance of DocuBits with a focus on individual learning. Second, we delve into the collaborative aspects, examining how DocuBits influence cooperative learning experiences for pairs of users engaged in a shared task with instructional documents.

For the within-subject study, participants undergo both individual and paired-user sessions. The individual study isolates DocuBits' features and examines their impact on cognitive load, immersion, learning performance, usability, and overall preference. The paired-
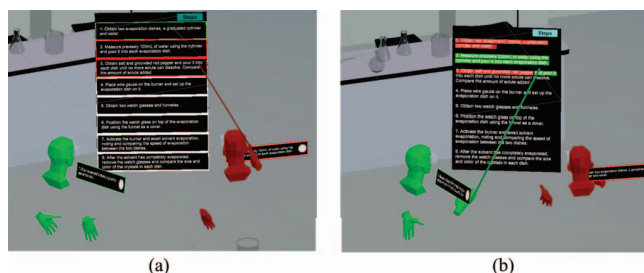

Figure 4: Users can allocate ownership of the DocuBits between multiple users by either (a) selecting pre-defined steps or (b) highlighting words.
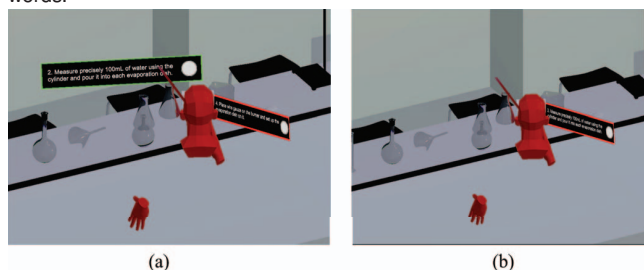

Figure 5: Users can change the ownership of an incomplete DocuBit by (a) selecting them and (b) the DocuBit will be stacked to the new owner's DocuBits.

user study extends the investigation, emphasizing collaboration-specific aspects. The overall hypotheses (OH1, OH2, OH3, OH4, OH5) were tested across both the single-user and paired-user studies, while collaboration-specific hypotheses (CH1, and CH2) were tested with the paired-user study only. The hypotheses are rooted in the expected user experience offered by DocuBits, which aligns with the feature design rationale detailed in Section 3.1 and Section 4.

Our hypotheses for the entire methodology were as follows:

- **(OH1)** *DocuBits reduce users' cognitive load during reading-while-doing.* We anticipate that the system's ability to allow users to focus on specific task elements will result in decreased cognitive demands.

- **(OH2)** *DocuBits enhance users' sense of immersion and presence in the virtual environment.* The tag-along-and-stick feature is expected to provide users with a more spatial sense when performing tasks.

- **(OH3)** *Users exhibit higher learning performance with DocuBits.* The reduced cognitive load and improved spatial awareness are anticipated to contribute to superior learning outcomes.

In the collaborative context, we introduce collaboration-specific hypotheses for the paired-user study:

- **(CH1)** *Balanced collaboration and teamwork will be enhanced with DocuBits.* DocuBits' capability to share each other's planned, ongoing, and completed/incomplete tasks will encourage equal task distribution and collaborative engagement.

- **(CH2)** *DocuBits foster higher co-presence and collaboration levels between paired users.* The ability to see each other's actions and actively solve task distribution issues is expected to enhance collaboration.

5.1 Implementation Setup

The experiment was set in a virtual chemistry lab environment built with Unity game engine. Each participant was equipped with an Oculus Quest 2 headset connected to a computer. The computer setup for each workstation included an Intel Xeon CPU and an NVIDIA GTX 1080 Ti graphics card. Participants comfortably sat

Figure 6: A top-down view of the virtual chemistry lab, with task locations color-coded: Lab 1 in yellow and Lab 2 in blue.

on swivel chairs during the study. For avatar movement within the virtual environment, participants had the option to turn around while seated to adjust their viewpoint. They also used controllers for aim-and-teleport actions to navigate the virtual space effectively. In the case of the paired-user study, we ensured that both participants were physically located in the same room. This arrangement prevented audio lag issues resulting from network latency, enabling seamless communication between participants.

Regarding text presentation, we adhered to established guidelines from prior work [15], employing a white Sans Serif Arial font for text display on a black background. This choice aimed to optimize text legibility and readability during the study. Text Mesh Pro was used for Unity implementation with a font size of 17.87 dmm based on previous work [30].

5.2 DocuBits for Single User

The primary objective of the first study is to evaluate the efficacy of DocuBits features when utilized by a single user, specifically focusing on their impact on individual learning performance and the overall VR experience during reading-while-doing tasks.

5.2.1 Participants

Eighteen participants (ten females, eight males) were recruited from a convenience sample of university students for the evaluation experiment (age range 22-34, $\mu = 27.78$, $\sigma = 4.88$). Ten participants wore glasses, one wore contact lenses, and the rest did not require vision correction. All our participants were proficient in English. Thirteen of the participants had previously experienced VR systems.

Study Procedure Steps
Instructions and Informed Consent Adjusting to the VR Headset and Environment Training with Methods
Experiment 1 Lab 1 with Method A Quiz Assessment Completion of Questionnaires Lab 1 Re-performance Assessment
Experiment 2 Lab 2 with Method B Quiz Assessment Completion of Questionnaires Lab 2 Re-performance Assessment
Post-Experiment Interview

Table 1: Study procedure followed by both Single-User and Paired-User studies, with methods presented in a counter-balanced order.

5.2.2 Procedure

Participants underwent a consent process, received training on the DocuBits method and a baseline monolithic document method, and completed two experiments with randomized assignments of these methods. In each experiment, participants first carried out a lab task using one of the document methods, followed by completing questionnaires about the experience. They then re-performed the same lab without any document method to assess learning retention. This process was repeated for a second lab task using the alternate document method.

Upon participant arrival, we detailed the study's protocol, secured consent via signed forms, and collected demographic data. Seated in swivel chairs, participants were acquainted with the Oculus Quest 2 and given time to familiarize themselves with the VR-rendered chemistry lab. Subsequently, they were given 10 minutes of training time to familiarize themselves with the baseline and DocuBits.

The study comprised two experiments inspired by distillation and crystallization labs. Each "lab" involves six tasks, such as retrieving and interacting with items, which required participants to move within the VR space. The tasks were fictional and designed not to benefit from prior chemistry knowledge which was clearly communicated to the participants. Task locations for each lab are illustrated in Figure 6. The order of labs was fixed, but the method the participants used (DocuBits or baseline) for each lab was randomized using a Latin Square design. We recorded from the start of the lab until verbal confirmation of completion.

Participants were directed to follow the provided document instructions, understanding that a post-lab quiz with seven questions would assess their task comprehension. The quiz focused on the sequence of tasks, observed results, and equipment locations. After completing the quiz, they filled out questionnaires on, in this order, presence (PQ [49]), immersion (IEQ [27]), cognitive load (NASA-RTLX [25]), and system usability (SUS [3]), and rated their preferences on a 5-point Likert scale. The presence questionnaire assesses the user's sense of 'being there' in the virtual space, while immersion evaluates engagement with the lab experiment. This approach aligns with prior works on collaborative learning in VR employing this questionnaire [17].

Participants then re-performed the lab experiment without document assistance to evaluate memory-based learning performance, with metrics including time and task accuracy. Following each experiment's completion, they repeated the process with the alternate document method. After both experiments, semi-structured interviews were conducted to collect detailed feedback. The entire procedure, lasting about 90 minutes, was video-recorded for analysis. Details are in Table 1.

5.2.3 Results

For analysis, we conducted a Wilcoxon-signed rank test for each measurement to determine the significance of our statistical results.

- **Completion Time:** DocuBits allowed participants to complete the tasks slightly faster than the baseline ($Md = 11.44min$, $IQR = 6.18$), baseline ($Md = 12.24min$, $IQR = 8.03$). However, there were no significant difference between the two methods ($Z = -0.806, p = 0.42$).

- **Cognitive Load:** DocuBits significantly reduced users' perceived cognitive load compared to the baseline in the single-user study. Users reported a lower cognitive load with DocuBits ($Md = 35.06$, $IQR = 22.05$) compared to the baseline ($Md = 54.62$, $IQR = 26.03$). This difference was statistically significant ($Z = -3.432, \mathbf{p} < \mathbf{0.001}$), highlighting the effectiveness of DocuBits in alleviating individual cognitive demands. This fulfills OH1:*DocuBits reduce users' cognitive load during reading-while-doing.*

- **Immersion and Presence:** DocuBits facilitated a higher sense of immersion ($Md = 4.13, IQR = 0.90$) compared to the baseline ($Md = 4.22, IQR = 1.41$), as well as presence ($Md = 3.53$, $IQR = 1.56$, baseline: $Md = 3.24, IQR = 1.27$). However, no significant difference was revealed ($Z = -0.816, p = 0.45$). While the scores are higher, it is hard to conclude that

WA Elements	Questions
Who-Presence	I was aware of my collaborator's presence in the virtual environment.
Who-Authorship	I was aware of what task my collaborator was responsible for.
What-Action	I was aware of what my collaborator was doing.
Where-Location	I was aware of where my collaborator was located.

Table 2: Collaboration questions based on workspace awareness from [24], rated on a 5-point Likert scale.

OH2:*DocuBits enhance users' sense of immersion and presence in the virtual environment* is true.

- **Learning Performance:** In terms of learning performance, there were no significant differences in quiz results in any of the two chemistry experiment's quiz between DocuBits (quiz1: $Md = 4.00$, $IQR = 3.00$, quiz2: $Md = 4.00$, $IQR = 2.25$) and the baseline (quiz1: $Md = 3.5$, $IQR = 2.25$, quiz2: $Md = 4.00$, $IQR = 2.25$) according to the Wilcoxon signed-rank test (quiz1: $p = 0.1$, quiz2: $p = 0.32$). On the other hand, for the re-performance of tasks without instructions, users had significantly higher scores (correct steps) with DocuBits($Md = 5.00$, $IQR = 3.00$), than the baseline ($Md = 3.00$, $IQR = 2.25$) with ($Z = -3.345$, $\mathbf{p < 0.001}$). The re-performance time was also significantly faster with DocuBits ($Md = 10.51min$, $IQR = 7.19$, baseline: $Md = 11.26min$, $IQR = 8.56$) ($Z = -2.069$, $\mathbf{p < 0.05}$). This indicates that DocuBits may have helped users learn the task better, aligning with OH3:*Users exhibit higher learning performance with DocuBits*.

- **Usability:** The SUS scores revealed that participants considered DocuBits($Md = 70.03$, $IQR = 23.49$) more usable than the baseline ($Md = 52.33$, $IQR = 25.52$). A statistical significance was revealed($Z = -3.432$, $\mathbf{p < 0.001}$).

- **Preference:** Participants reported a significantly higher preference towards DocuBits ($Md = 4.00$, $IQR = 2.00$) compared to the baseline ($Md = 2.50$, $IQR = 1.00$, $Z = -3.342$, $\mathbf{p < 0.001}$).

5.3 DocuBits for Paired Users

In this experiment, we designed the study to require two users to work together to complete a task. Our focus was to determine the impact of having an interaction framework that allows shareable, trackable, interactive document fragments between collaborators.

5.3.1 Participants

We recruited 14 pairs ($N = 28$, 18 male, 10 female) from a convenience sample of university students (age range 22-34, $\mu = 25.7$, $\sigma = 3.50$). Among the participants, 21 had prior experience with VR systems, and all were proficient in English. Participant pairs were selected such that they were already acquainted with each other, ensuring a minimum level of communication confusion due to social unfamiliarity. No participants were re-recruited from Section 5.2.

5.3.2 Procedure

The paired-user study also follows the procedure listed in Table 1. Here, participants were seated in swivel chairs at individual workstations in the same room. They were given time to adjust themselves to the Oculus Quest 2 and the virtual chemistry lab environment, followed by a 10-minute training to familiarize themselves with both DocuBits and the baseline.

The labs, inspired by distillation and crystallization labs, comprised eight steps distinct from those in the single-user study. Pairs proceeded through each lab using one of the two methods, with method presentation order randomized via Latin Square. Participants had complete autonomy in assigning tasks within their pair, and we recorded the time taken to complete the experiment, including task distribution, culminating when they verbally confirmed task completion. To measure task distribution among participants, we counted for each participant (i) how many tasks they finished on their own, and (ii) how many tasks were done collaboratively (completed together or taken over by another participant).

Subsequently, participants individually completed a post-trial quiz comprising seven questions without discussion. Following the quiz, they completed the same set of questionnaires as in the single-user study, assessing immersion, presence, cognitive load, usability, and preference. Additionally, we inquired about their perceived social presence in collaborative learning [33], adapting the items to our study's context, and gathered data on collaboration-specific factors based on the workspace awareness framework [24], detailed in Table 2. Participants then engaged in a re-performance assessment, collaboratively executing the experiment steps without written instructions.

Post-experiment, we conducted semi-structured interviews with the pairs concurrently to capture comprehensive feedback and reflections. The entire procedure spanned approximately 120 minutes, with all sessions video-recorded for in-depth analysis

5.3.3 Results

Similar to Section 5.2.3, a Wilcoxon signed-rank test was conducted for each measurement.

- **Completion Time:** DocuBits exhibited a significantly shorter completion time ($Md = 15.42min$, $IQR = 8.81$) compared to the baseline ($Md = 17.39min$, $IQR = 11.17$) ($Z = -3.158$, $\mathbf{p < 0.01}$).

- **Cognitive Load:** Users reported a reduced cognitive load with DocuBits ($Md = 34.64$, $IQR = 24.00$) in contrast to the baseline ($Md = 57.48$, $IQR = 24.11$). This statistically significant difference ($Z = -3.552$, $\mathbf{p < 0.001}$) underscores the effectiveness of DocuBits in alleviating user cognitive demands (OH1).

- **Immersion and Presence:** DocuBits facilitated a heightened sense of immersion ($Md = 6.13$, $IQR = 2.36$) compared to the baseline ($Md = 4.27$, $IQR = 1.47$) and presence ($Md = 4.72$, $IQR = 2.23$, baseline: $Md = 3.28$, $IQR = 1.18$). This significant difference ($Z = -4.213$, $\mathbf{p < 0.001}$) suggests that users felt more engaged and immersed when using DocuBits (OH2).

- **Task Distribution:** Task distribution was assessed by comparing the difference in the number of tasks performed by an individual within a pair. DocuBits ($Md = 1.00$, $IQR = 1.00$) demonstrated a more equal contribution among users compared to the baseline ($Md = 1.00$, $IQR = 2.00$). However, no significant difference was revealed ($Z = -1.811$, $p = 0.70$). On the other hand, DocuBits exhibited a higher count of tasks performed collaboratively ($Md = 2.00$, $IQR = 0.25$) compared to the baseline ($Md = 1.00$, $IQR = 2.00$). The results were statistically different ($Z = -2.739$, $p = 0.006$). Hence, while there are implications of equal task distribution, we can only partially support CH1: *Balanced collaboration and teamwork will be enhanced with DocuBits.*.

- **Social Presence and Collaboration:** Participants reported a higher perceived social presence with DocuBits ($Md = 4.55$, $IQR = 1.02$) compared to the baseline ($Md = 3.00$, $IQR = 1.97$). DocuBits excelled in promoting a higher level of collaboration among users ($Md = 4.14$, $IQR = 0.93$) compared to the baseline ($Md = 2.49$, $IQR = 1.76$). The observed difference is statistically significant ($Z = -4.463$, $\mathbf{p < 0.001}$), indicating
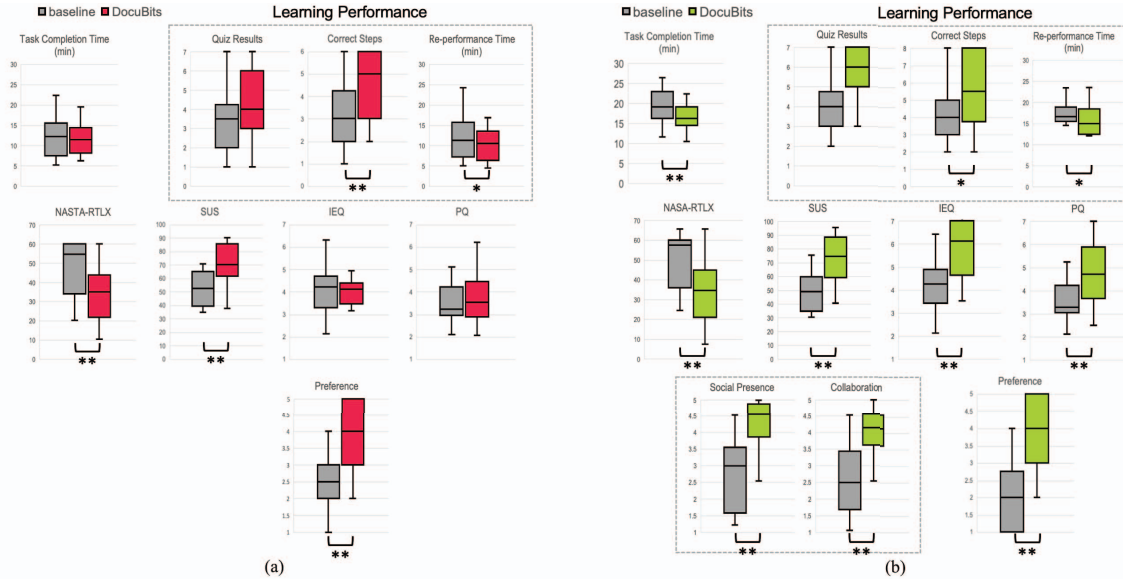
Figure 7: Statistical results for (a) the single-user study and (b) the paired-user study. In the single-user study, statistically significant results were observed for correct steps in re-performing the task, workload (NASA-RTLX), usability (SUS), and preference. In the paired-user study, statistically significant results were observed for task completion time, correct steps, re-performance time, workload (NASA-RTLX), usability (SUS), immersion (IEQ), presence (PQ), social presence, collaboration (questions described in Table. 2), and preference. Asterisk (*) indicates a statistically significant difference between conditions: $p < 0.05 (*)$; $p < 0.001 (**)$.

a strong preference for collaborative tasks using DocuBits. This supports CH2: *DocuBits foster higher co-presence and collaboration between paired users*.

- **Learning Performance:** In line with the single-user study, the two quiz results were comparable between DocuBits (quiz1: $Md = 6.00, IQR = 2.00$, quiz2: $Md = 5.00, IQR = 1.75$) and the baseline (quiz1: $Md = 4.00, IQR = 1.75$, quiz2: $Md = 6.00, IQR = 2.00$) (quiz1: $p = 0.46$, quiz2: $p = 0.1$). However, users took less time to recreate the lab experiment with DocuBits ($Md = 13.43min, IQR = 6.71$) than with the baseline ($Md = 16.02min, IQR = 6.59$). This time-saving benefit with DocuBits was statistically significant ($Z = -2.417, \mathbf{p < 0.05}$). Additionally, users demonstrated a significant difference in the number of correct steps, with DocuBits scoring higher ($Md = 5.52, IQR = 4.00$) than the baseline ($Md = 4.50, IQR = 3.31, Z = -2.622, \mathbf{p < 0.01}$). This supports the fact that DocuBits elicit users to have higher learning performances(OH3).

- **Usability:** Participants rated DocuBits higher in usability ($Md = 74.67, IQR = 29.26$) compared to the baseline ($Md = 49.11, IQR = 24.95$). This significant difference ($Z = -4.031, \mathbf{p < 0.001}$) affirms the enhanced usability of DocuBits in virtual collaborative environments.

- **Preference:** DocuBits had a higher preference score ($Md = 4.00, IQR = 2.00$) than the baseline ($Md = 2.00, IQR = 1.75$), which was statistically significant ($Z = -4.266, \mathbf{p < 0.001}$).

Here, our approach, DocuBits, consistently outperformed the baseline across all metrics, establishing its effectiveness in collaborative learning scenarios.

6 DISCUSSION

In this section, we discuss the results from the two evaluation studies presented in Section 5. We focus on how the results support or do not support our hypotheses, using quotes from participants from the post-experiment interviews for illustration. Participants are numbered continuously, with P1-P18 from the single-user study and P19-P48 from the paired-user study.

6.1 DocuBits for Facilitating Collaboration

Our result analysis revealed DocuBits offered beneficial results in both the single-user study (OH1, OH3) and the paired-user study (OH1, OH2, OH3, CH1, CH2). This indicates that DocuBits could be a valuable tool in either setting, with the added benefits of greater immersion and collaboration in the paired setting. Particularly, the higher number of tasks completed collaboratively by paired participants using DocuBits supports the notion of balanced collaboration (CH1). This trend demonstrates DocuBits' potential to foster an equitable task distribution among users.

Participants emphasized that the interface facilitated collaborative decision-making sessions, aided visually by DocuBits. The color-coded interface allowed for immediate tracking of the distribution of work between users. P12 highlighted, "*The process of selecting which task I am going to perform made me more conscious of how much work I was doing. I wanted to make sure my partner was not doing all the work for me.*" Even during the re-performance test, participants exhibited a better recollection of who had performed which task with DocuBits, attributing it to the specific procedural guidance provided. Similarly, P7 expressed, "*I was constantly aware of the workload balance, thanks to the visual cues from DocuBits.*" Pair P43 and P44 noted the shared features of DocuBits facilitated their cooperative work. P43 remarked, "*That floating DocuBit with an 'X' on it really grabbed my attention. I just had to jump in and help my partner out so we could get it to stay put!*"

6.2 Preferred Method for Creating DocuBits

In the observation of participants' preferences for creating DocuBits, a prominent trend emerged, showcasing a clear inclination towards the stepped method. This preference was particularly pronounced when participants engaged in individual tasks. P19 reflected, "*The stepped method provided a structured and organized approach. It felt like a natural way to break down complex instructions into manageable steps, especially when I was working through them on my own. It helped maintain a clear sequence and aided in a more systematic execution of tasks.*"

On the other hand, some participants (P20, P21, P24, P25, P31)

316

mentioned that their preference for the stepped method was due to convenience, but they acknowledged the potential need for the highlighting method in more complex tasks. Particularly, the user pair P24 and P25 collectively recalled, "*the stepped method was easier with the instructions we had, but in cases where a single step involves multiple substeps, the highlight method would come in handy.*" P31 suggested the usefulness of being able to switch between methods, saying, "*Ultimately, I would prefer if I could do both combined. Sometimes I wanted to make minor adjustments to the pre-fragmented steps and I don't want to spend too much time highlighting every step.*"

This inclination towards a stepped approach underscores the importance of tailoring the DocuBits interface to individual work styles, ensuring that the method of creating these instructional cues aligns with user preferences. An intuitive and adaptable creation method could significantly contribute to the user's efficiency and comfort, ultimately enhancing the overall user experience.

6.3 Need for Ownership Support

Participants articulated a compelling need for enhanced ownership support, a crucial aspect when collaborating on tasks within the DocuBits framework. The majority of the participant pairs in the paired user study pointed out how they liked the idea of transferring ownership of a DocuBit. A participant pair, P19, P20 stated, "The smooth transitions of ownership not only enhances the collaborative experience but also allowed more freedom in collaboration styles and preferences."

A noteworthy perspective surfaced, with participants expressing a desire to assign the same task to multiple individuals and have it automatically marked as complete once any one of them finished. As P34 emphasized, "*In collaborative settings, there are often times you wish to work on the same task with your collaborator or allow whoever finishes the previous task first to take on the task. It would be more practical if there could be a way we can have access to a DocuBit at the same time. The current version only allows one owner per DocuBit.*" This participant insight underscores the significance of refining collaborative features within DocuBits to accommodate diverse task distribution scenarios.

6.4 Observed Collaboration Styles

Diverse collaboration styles were observed among participants, revealing intriguing insights into task allocation strategies within the DocuBits framework. Some users used a strategy of pre-placing all DocuBits before initiating the procedural task, reflecting a meticulous planning approach. P38 elucidated, "*I preferred setting everything up first. It provided a comprehensive visualization of the entire process before diving into the actual tasks. This strategic planning helped me allocate tasks more effectively.*"

Additionally, users displayed a proclivity for dividing tasks evenly or allocating specific types of tasks to different users, showcasing the flexibility and adaptability of DocuBits to varied collaboration styles. For instance, participants chose to exclusively handle apparatus-related tasks as a specialized allocation strategy. These insights highlight the diverse ways users approach collaborative work within the DocuBits environment, emphasizing the need for a flexible and intuitive tool that accommodates these varied preferences.

6.5 Effect of Texts on Collaboration

Participants shed light on the nuanced impact of text-based instructions on collaboration, revealing an interesting dynamic within the DocuBits framework. Some participants expressed a need for integration with demonstration videos, envisioning a more comprehensive guidance approach. P44 suggested, "*Having the option to seamlessly switch between text and video instructions could significantly enhance task distribution and performance. It caters to different learning preferences and ensures a more versatile collaboration experience.*"

This participant insight unveils a potential avenue for future work, pointing towards the integration of textual and visual guidance within DocuBits. Such a feature could provide users with versatile options for task execution and distribution, further enhancing the collaborative and instructional capabilities of the tool. As DocuBits evolves, addressing these nuanced preferences for instructional content could substantially contribute to the framework's adaptability and user-centric design.

7 LIMITATIONS AND FUTURE WORK

Our method's efficacy in comparison with documents featuring scrolls and pages warrants exploration. The frequency of feature usage is anticipated to vary based on task characteristics. For instance, users might utilize the highlight feature more for excessively long steps or when tasks are perceived as intricate.

The current implementation focuses on collaboration between two users. Scaling the tool for more than two users necessitates refined design considerations. Improved interaction methods for claiming DocuBit ownership should address potential confusion regarding task attribution. Integration with the spatial map could assist in tracking task locations and current user engagement. Additionally, enhancing color-coded ownership indicators with user-specific icons can enhance clarity.

As for future work, we envision advancing DocuBits into a seamless pipeline. Future work will involve:

- **Automatic Instructional Document Integration:** Development of an automatic pipeline to scan and import instructional documents into DocuBits.
- **Automated Performance Evaluation:** Implementation of an automated performance evaluation mechanism to assess task completion.
- **Instructor-Side Connectivity:** Sharing task performance results with the instructor side for real-time analysis of student performance and identification of areas requiring assistance.

Our current prototype requires clear numbering or manual entry of instructions into DocuBits. Future implementations aim to extend this by facilitating the easy selection of typical separators. We foresee establishing labeled "anchor points" during authoring to automate DocuBit placement. Expansion beyond two users is anticipated, with a proposed special "birdseye" view for instructors in classroom scenarios. Furthermore, we envision extending DocuBits to Mixed Reality, utilizing MR anchors in real space instead of VR coordinates.

8 CONCLUSION

We introduced DocuBits, a novel method for decomposing monolithic documents into portable interactive elements. Informed by insights from a needs-finding study with 7 experienced educators, we delineated key design considerations for VR interfaces handling text instructions. DocuBits facilitated procedural task completion by (i) fragmenting text instructions, (ii) anchoring them to task locations, and (iii) enabling progress monitoring and sharing among users. In a comprehensive two-fold user study—first with a single user and then with paired users—we established that DocuBits enhanced learning performance in task recall and execution, while also positively impacting perceived workload, presence, immersion, and usability.

We further discussed how DocuBits fostered collaboration, user strategies in creating DocuBits, necessary support for multi-user DocuBits assignments, observed collaboration styles between paired users, and the potential influence of text instructions on collaboration dynamics. We anticipated DocuBits' extension for supporting automated performance evaluation in VR training applications and its integration into AR/MR contexts.

REFERENCES

[1] P. Adami, P. B. Rodrigues, P. J. Woods, B. Becerik-Gerber, L. Soibelman, Y. Copur-Gencturk, and G. Lucas. Effectiveness of vr-based training on improving construction workers' knowledge, skills, and safety behavior in robotic teleoperation. *Advanced Engineering Informatics*, 50:101431, 2021.

[2] S. Bailey, C. I. Johnson, B. L. Schroeder, and M. D. Marraffino. Using virtual reality for training maintenance procedures. In *Proceedings of the Interservice/Industry Training, Simulation and Education Conference*, 2017.

[3] A. Bangor, P. T. Kortum, and J. T. Miller. An empirical evaluation of the system usability scale. *Intl. Journal of Human–Computer Interaction*, 24(6):574–594, 2008.

[4] I. Becker, V. Toivonen, S.-P. Leino, et al. Using virtual reality in designing the assembly process of a car. In *DS 68-9: Proceedings of the 18th International Conference on Engineering Design (ICED 11), Impacting Society through Engineering Design, Vol. 9: Design Methods and Tools pt. 1, Lyngby/Copenhagen, Denmark, 15.-19.08. 2011*, pp. 91–100, 2011.

[5] G. Chen, W. Cheng, T.-W. Chang, X. Zheng, and R. Huang. A comparison of reading comprehension across paper, computer screens, and tablets: Does tablet familiarity matter? *Journal of Computers in Education*, 1(2):213–225, 2014.

[6] J. Chen, P. S. Pyla, and D. A. Bowman. Testbed evaluation of navigation and text display techniques in an information-rich virtual environment. In *IEEE Virtual Reality 2004*, pp. 181–289. IEEE, 2004.

[7] E. Childs, F. Mohammad, L. Stevens, H. Burbelo, A. Awoke, N. Rewkowski, and D. Manocha. An overview of enhancing distance learning through emerging augmented and virtual reality technologies. *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–15, 2023. doi: 10.1109/TVCG.2023.3264577

[8] P. T. Chua, R. Crivella, B. Daly, N. Hu, R. Schaaf, D. Ventura, T. Camill, J. Hodgins, and R. Pausch. Training for physical tasks in virtual environments: Tai chi. In *IEEE Virtual Reality, 2003. Proceedings.*, pp. 87–94. IEEE, 2003.

[9] C. Connell, L. Bayliss, and W. Farmer. Effects of ebook readers and tablet computers on reading comprehension. *International Journal of Instructional Media*, 39(2), 2012.

[10] P. Corporation. Introduction photon engine, 2023.

[11] P. Delgado, C. Vargas, R. Ackerman, and L. Salmerón. Don't throw away your printed books: A meta-analysis on the effects of reading media on reading comprehension. *Educational Research Review*, 25:23–38, 2018.

[12] E. Dieterle. Multi-user virtual environments for teaching and learning. In *Encyclopedia of Multimedia Technology and Networking, Second Edition*, pp. 1033–1041. IGI Global, 2009.

[13] A. Dillon. Reading from paper versus screens: A critical review of the empirical literature. *Ergonomics*, 35(10):1297–1326, 1992.

[14] A. Dillon, J. Richardson, and C. McKnight. The effects of display size and text splitting on reading lengthy text from screen. *Behaviour & Information Technology*, 9(3):215–227, 1990.

[15] T. Dingler, K. Kunze, and B. Outram. Vr reading uis: Assessing text parameters for reading in vr. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–6, 2018.

[16] E. Dittrich, S. Brandenburg, and B. Beckmann-Dobrev. Legibility of letters in reality, 2d and 3d projection. In *International Conference on Virtual, Augmented and Mixed Reality*, pp. 149–158. Springer, 2013.

[17] T. Drey, P. Albus, S. der Kinderen, M. Milo, T. Segschneider, L. Chanzab, M. Rietzler, T. Seufert, and E. Rukzio. Towards collaborative learning in virtual reality: A comparison of co-located symmetric and asymmetric pair-learning. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pp. 1–19, 2022.

[18] J. Gabel, M. Ludwig, and F. Steinicke. Immersive reading: Comparison of performance and user experience for reading long texts in virtual reality. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, pp. 1–8, 2023.

[19] X. Gao, R. Gong, T. Shu, X. Xie, S. Wang, and S.-C. Zhu. Vrkitchen: An interactive 3d environment for learning real life cooking tasks. 2019.

[20] M. García García, J. M. Arenas Reina, A. S. Lite, and M. Á. Sebastián Pérez. Simulation of assembly processes with technical of virtual reality. In *AIP Conference Proceedings*, vol. 1181, pp. 662–669. American Institute of Physics, 2009.

[21] N. Gavish, T. Gutiérrez, S. Webel, J. Rodríguez, M. Peveri, U. Bockholt, and F. Tecchia. Evaluating virtual reality and augmented reality training for industrial maintenance and assembly tasks. *Interactive Learning Environments*, 23(6):778–798, 2015.

[22] J. Georgiou, K. Dimitropoulos, and A. Manitsaris. A virtual reality laboratory for distance education in chemistry. *International Journal of Social Sciences*, 2(1):34–41, 2007.

[23] J. D. Gould, L. Alfaro, R. Finn, B. Haupt, and A. Minuto. Reading from crt displays can be as fast as reading from paper. *Human factors*, 29(5):497–517, 1987.

[24] C. Gutwin and S. Greenberg. A descriptive framework of workspace awareness for real-time groupware. *Computer Supported Cooperative Work (CSCW)*, 11:411–446, 2002.

[25] S. G. Hart. Nasa-task load index (nasa-tlx); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, vol. 50, pp. 904–908. Sage publications Sage CA: Los Angeles, CA, 2006.

[26] J. Jankowski, K. Samp, I. Irzynska, M. Jozwowicz, and S. Decker. Integrating text with video and 3d graphics: The effects of text drawing styles on text readability. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1321–1330, 2010.

[27] C. Jennett, A. L. Cox, P. Cairns, S. Dhoparee, A. Epps, T. Tijs, and A. Walton. Measuring and defining the experience of immersion in games. *International journal of human-computer studies*, 66(9):641–661, 2008.

[28] Q. Jin, Y. Liu, R. Sun, C. Chen, P. Zhou, B. Han, F. Qian, and S. Yarosh. Collaborative online learning with vr video: Roles of collaborative tools and shared video control. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pp. 1–18, 2023.

[29] S. Kobayashi, K. Kanari, and M. Sato. An examination of view-settings for long texts in vr reading. In *ACM SIGGRAPH 2021 Posters*, pp. 1–2. 2021.

[30] T. Kojić, D. Ali, R. Greinacher, S. Möller, and J.-N. Voigt-Antons. User experience of reading in virtual reality—finding values for text distance, size and contrast. In *2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX)*, pp. 1–6. IEEE, 2020.

[31] G. Lee, J. Healey, and D. Manocha. Vrdoc: Gaze-based interactions for vr reading experience. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 787–796. IEEE, 2022.

[32] J. Lee, M. Kim, and J. Kim. Rolevr: Multi-experience in immersive virtual reality between co-located hmd and non-hmd users. *Multimedia Tools and Applications*, 79:979–1005, 2020.

[33] G.-Y. Lin. Social presence questionnaire of online collaborative learning: Development and validity. *Association for Educational Communications and Technology*, 2004.

[34] R. B. Loftin and P. Kenney. Training the hubble space telescope flight team. *IEEE Computer Graphics and Applications*, 15(5):31–37, 1995.

[35] T. D. Parsons, T. Bowerly, J. G. Buckwalter, and A. A. Rizzo. A controlled clinical comparison of attention performance in children with adhd in a virtual reality classroom compared to standard neuropsychological methods. *Child neuropsychology*, 13(4):363–381, 2007.

[36] K. Patel, J. N. Bailenson, S. Hack-Jung, R. Diankov, and R. Bajcsy. The effects of fully immersive virtual reality on the learning of physical tasks. In *Proceedings of the 9th Annual International Workshop on Presence, Ohio, USA*, pp. 87–94, 2006.

[37] I. Radu, V. Hv, and B. Schneider. Unequal impacts of augmented reality on learning and collaboration during robot programming with peers. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW3):1–23, 2021.

[38] P.-L. P. Rau, J. Zheng, Z. Guo, and J. Li. Speed reading on virtual reality and augmented reality. *Computers & Education*, 125:240–245, 2018.

[39] D. Schott, P. Saalfeld, G. Schmidt, F. Joeres, C. Boedecker, F. Huettl, H. Lang, T. Huber, B. Preim, and C. Hansen. A vr/ar environment for multi-user liver anatomy education. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*, pp. 296–305. IEEE, 2021.

[40] A. L. Simeone, M. Speicher, A. Molnar, A. Wilde, and F. Daiber. Live: The human role in learning in immersive virtual environments. In *Symposium on spatial user interaction*, pp. 1–11, 2019.

[41] S. Stansfield, D. Shawver, and A. Sobel. Biosimmer: A virtual reality simulator for training first responders in a bw scenario. In *Proceedings of the 1998 Scientific Conference on Chemical and Biological Defense Research*, 1999.

[42] J. D. Stokes, A. Rizzo, J. J. Geng, and J. B. Schweitzer. Measuring attentional distraction in children with adhd using virtual reality technology with eye-tracking. *Frontiers in virtual reality*, 3:855895, 2022.

[43] D. L. Tate, L. Sibert, and T. King. Virtual environments for shipboard firefighting training. In *Proceedings of IEEE 1997 Annual International Symposium on Virtual Reality*, pp. 61–68. IEEE, 1997.

[44] M. Thompson, H. Olivas-Holguin, A. Wang, J. Fan, K. Pan, D. Vargas, and J. Gerr. Rules, roles, and resources: Strategies to promote collaboration in virtual reality contexts. In *Workshop Position Paper for CHI 2018*, 2018.

[45] B. Thoravi Kumaravel, F. Anderson, G. Fitzmaurice, B. Hartmann, and T. Grossman. Loki: Facilitating remote instruction of physical tasks using bi-directional mixed-reality telepresence. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, pp. 161–174, 2019.

[46] E. Trondsen and K. Vickery. Learning on demand. *Journal of Knowledge Management*, 1(3):169–180, 1997.

[47] C. Uz-Bilgin, M. Thompson, and M. Anteneh. Exploring how role and background influence through analysis of spatial dialogue in collaborative problem-solving games. *Journal of Science Education and Technology*, 29:813–826, 2020.

[48] C. Wei, D. Yu, and T. Dingler. Reading on 3d surfaces in virtual environments. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 721–728. IEEE, 2020.

[49] B. G. Witmer and M. J. Singer. Measuring presence in virtual environments: A presence questionnaire. *Presence*, 7(3):225–240, 1998.